

现代数学译丛 28

利己主义的数学解析

〔奥地利〕K.Sigmund 著

徐金亚 杨 静 汪 芳 译

陆征一 校



科 学 出 版 社

现代数学译丛 28

利己主义的数学解析

〔奥地利〕 K. Sigmund 著
徐金亚 杨 静 汪 芳 译
陆征一 校

科 学 出 版 社
北 京

图字：01-2016-9609

内 容 简 介

本书著者卡尔·西格蒙德(Karl Sigmund)为国际著名数学家和生物数学家，奥地利科学院院士. 在动力系统、种群动力学、进化博弈理论等方面做出了一批开创性的工作.

本书以合作，背叛为基本概念. 界定利己，即趋利避害. 然后给出善良以及各阶好人的定义. 在首先描述了亲缘、邻里关系等动物特性以外，重申亚当·斯密的观点：利己作为无形之手，将个体的选择转化为群体的整体利益. 从捐赠博弈开始，分别介绍了囚徒困境博弈、最后通牒博弈、独裁者博弈、信任博弈、重复囚徒困境博弈等诸多重要的博弈类型.

本书通过复制方程对主要的博弈进行数学表述，并利用动力系统的一些基本方法，详细分析了对应于具有线性收益与非线性收益的复制方程，给出了群体在不同状态下的动力学行为，揭示其进化与发展趋势.

本书涉及哲学、社会学、生物学、心理学、神经学以及数学等众多领域，是一部以概念、集直观感性和光辉理性为一体的饕餮思辨之大餐.

图书在版编目(CIP)数据

利己主义的数学解析/ (奥) 卡尔·西格蒙德 (K. Sigmund) 著: 徐金亚, 杨静, 汪芳译. —北京: 科学出版社, 2017. 8
(现代数学译丛; 28)
书名原文: The Calculus of Selfishness
ISBN 978-7-03-054075-1

Ⅰ.①利… Ⅱ. ①卡… ②徐… ③杨… ④汪… Ⅲ. ①博弈论 Ⅳ. ①O225

中国版本图书馆 CIP 数据核字 (2017) 第 185123 号

责任编辑: 陈玉琢 / 责任校对: 彭珍珍
责任印制: 张 伟 / 封面设计: 陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

北京东华虎彩印刷有限公司 印刷
科学出版社发行 各地新华书店经销
*
2017 年 8 月第 一 版 开本: 720 × 1000 1/16
2017 年 8 月第一次印刷 印张: 11 1/2
字数: 230 000

定价: 68. 00 元
(如有印装质量问题, 我社负责调换)

校 译 者 序

本书根据利己、合作与背叛等基本概念, 给出了善良以及直接或间接回报 (报答合作、报复背叛) 等概念, 在极少的公理化假设下, 借助复制动力系统, 研究了由理性支配的利己如何走向合作.

著者在前言中指出, 该书可以作为应用科学的本科生教材. 但实际上, 其所需的数学知识特别是微分方程定性分析的内容超出了非数学专业本科生的范围. 对少数学习了基本定性理论: 局部稳定分析, Pioncaré-Bendison 定理, Liapunov 稳定性等知识的本科数学专业学生在接受基本哲学观点的前提下可以阅读. 就校译者观点, 本书是一部具有丰富背景及数学意味的专著, 适合研究生及科研工作者阅读.

校译者对于书名的翻译有两个选择, 或是《自私的计算》. 选择《利己主义的数学解析》可能更能准确地表达著者的原意. 除了几处必须之外, 本书均选择利己而非自私, 主要希望去除感情色彩以本性而论.

我们与著者保持了长期的学术联系. 本书的翻译和出版是著者和校译者的共同愿望. 通过学习和翻译本书, 杨静利用Lotka-Volterra 系统和复制系统的等价关系并借助 Kolmogorov 系统得到了一个论证高维Lotka-Volterra 系统极限环存在的新技巧. 汪芳从著者论证中心存在的工作中受到启发, 给出了复制系统存在中心及正平衡关于全局稳定的充要条件. 管华明将无条件策略与二阶好人的博弈推广到了无条件策略与三阶好人博弈, 得到了无条件策略与最顶尖八大策略的完整动力学性质.

本书的翻译分工如下: 徐金亚负责第 1, 第 4 和第 7 章, 杨静负责第 2 和第 5 章, 汪芳负责第 3 和第 6 章, 陆征一负责审校. 在历时两年多的学习以及其后的翻译和校对中, 我们得到了众多同仁一次至多次的帮助. 感谢路过和参与讨论班的杜亚佳, 管华明, 何碧, 刘春燕, 刘静, 马思碧, 王照琨, 吴文文, 杨德牛等. 特别感谢杨鸣, 他自始至终参与了讨论, 并在文字, 图形以及参考文献的修改, 处理和编排等方面给予大力帮助, 他的贡献使译本大大增色. 感谢王维民教授在翻译期间给予的雪中送炭和锦上添花的指导.

我们希望该书的中文版可以忠实地将个人主义方法论, 合作的进化以及复制动力学完美结合的最新和最高水平的成果呈现给读者.

校译者
2017 年 4 月

著者前言

本书并非面向富有心计的利己 (自私) 者. 实际上, 我们的兴趣在于社会生活中随处可见的讨价还价其背后所隐藏的逻辑. 本书面向的读者群主要是经济学、心理学、进化生物学等专业的本科生, 而社会困境的简单模型是本书的主题.

正如书名中 "数学解析" 这一词所暗示的, 你需要具备一些初步的数学知识. 另外, 书中以容易接受的方式循序渐进地介绍了所需的博弈论知识. 至于书名中的 "利己" 一词, 并非指盲目的贪婪, 而是意味着理性的利己, 正如托克维尔 (Tocqueville) 所说, "美国人喜欢解释生活中的任何行为; …… 他们自鸣得意地宣扬, 理性的利己如何最终促进了相互帮助." 这种自鸣得意可能是合理的; 但是理论学家不能分享这一点. 因为许多人难以理解为什么源于利己的动机却能促成人们为共同利益而进行合作.

在 2000 年《科学》杂志的新年号上, 主编们将 "合作的进化" 列为新世纪最具挑战性的十大问题之一. 本书并非声称解决了这个问题. 在该领域工作了二十多年, 我深知其发展非常迅速, 即便是局限于人类社会的合作这一生物学合作问题中的细小领域, 也难以仅用一本书来呈现其丰富的内容.

本书并不试图解决所有问题, 而只讨论利己个体之间的回报, 并探究少许基本的交互问题. 研究方法基于进化的准则: 越成功的策略在群体内越流行. 我们不考虑家庭纽带、邻里关系、个体及文化差异. 在开篇就指出这一限定最好不过了. 我不希望本书误导大家产生这样的印象: 家庭纽带、社区关系或个人因素等等在合作的进化中无关紧要, 而只剩下利己. 正如理论物理学家撰写关于万有引力的专著, 并非因此表明宇宙中其他作用力就可以忽略. 据有关分析, 当前经济发展的趋势似乎使经济交互正脱离裙带关系, 狭隘主义以及面对面交易的局限, 而扩大至融合良好的社会中的任意陌生人之间.

引言部分 (通篇不含任何公式) 描述了一些最基本的社会困境. 历代思想家都着迷于利己主义与利他主义行为问题的讨论. 但是, 直至最近才引入了正式的模型以及博弈实验. 自从罗伯特 · 特里弗斯 (Robert Trivers) 在回报中引入了进化的手段, 囚徒困境才成为研究该问题的一种模式生物 (model organism). 而其他博弈如最后通牒博弈的研究很快就迎头赶上. 最振奋人心的是关于理论进展和实验进展之间的紧密联系.

第 2 章独立地介绍进化博弈论, 结合经典博弈论的核心概念, 譬如纳什均衡、风险占优等, 重点介绍确定性动力学和随机过程理论.

第 3 章详细讨论重复博弈, 如囚徒困境博弈和雪堆博弈, 借此研究了反复博弈在两个参与者之间的直接回报. 特别地, 一些简单的策略, 譬如, 基于前一回合结果的*以牙还牙* (Tit for Tat) 策略, 由有限自动机实现的*悔悟的以牙还牙* (Contrite Tit for Tat) 策略等, 呈现了丰富的行为.

第 4 章专门讨论了间接回报. 参与者至多相遇一次, 但他们知道对手的历史行为. 其中引入了诸如道德判断、声誉关注等因素. 对于基于第三方之间交互的评价之上的策略, 考虑了不受剥削影响的合作类型, 因为此类合作只针对合作者.

第 5 章研究了最后通牒博弈和信任博弈. 这类博弈用于探讨公平与信任问题, 此外的有益之处在于提供了一个分析正向和负向激励作用的框架. 声誉对于合作的促成再次起到了至关重要的作用.

第 6 章的讨论由两个参与者的博弈转向大群体内的博弈. 在所谓的公共利益博弈中, 可以通过奖惩机制对背叛行为进行压制. 通过适当的激励机制可以在合作者和背叛者的混合群体中建立回报. 一个有趣的工作讨论了引入可以自愿参与而非必须参与团队合作的情形. 如果可选择是否参与, 那么强制更容易出现.

最后, 篇幅较短的第 7 章简要讨论之前未予考虑的因素, 譬如裙带关系, 本位主义以及群体选择等. 不用说, 本书很大程度上要归功于我的同事们, 他们中的许多人通读了全书并提出了宝贵的建议. 特别地, 我要感谢 Christoph Hauert, Arne Traulsen, Hannelore De Silva (原名 Brandt), Hisashi Ohtsuki, Satoshi Uchida, Ulf Dieckmann, Tatsuya Sasaki, Simon Levin, Ross Cressman, Yoh Iwasa, Silvia De Monte, Christoph Pflügl, Christian Hilbe, Steve Frank, Simon Gächter, Benedikt Hermann, Dirk Semmann, Manfred Milinski 以及 Josef Hofbauer. 这里面, 我要特别感谢马丁 · 诺瓦克 (Martin Nowak), 没有他, 本书不可能完成.

目　录

第1章　引言：社会陷阱与简单博弈

1.1　社会性动物

亚里士多德将人类和蚂蚁、蜜蜂等生物归类为社会性动物. 此后, 大量作者将人类社会的城邦和国家与蜂巢和蚁丘作比较, 如伯纳德 · 曼德维尔 (Bernard de Mandeville) 在三百多年前出版的《蜜蜂寓言》(*The Fable of the Bees*).

今天, 我们知道人类社会与昆虫王国差别并不太大. 群居昆虫所表现的高度惊人的合作性本质上是源于蚁丘和蜂巢内稳固的家庭纽带. 相比而言, 人类会经常与没有亲缘关系的伙伴合作.

与近亲之间的合作称为亲缘选择 (kin selection). 愿意养育后代的基因显然有利于自身的传播. 愿意帮助兄弟姐妹的基因同样也有利于自身的传播, 只是这种传播不是直接通过自己的嫡系后代而是通过同胞兄弟姐妹的后代. 实际上, 近亲极有可能携带与自己相同的基因. 在蜂巢里, 所有的工蜂都是姐妹, 而蜂王则是它们的母亲. 蜂王可能有多个配偶, 于是工蜂的平均亲缘相关度随之降低. 亲缘选择理论较为复杂且尚有争议. 然而家庭关系对于解释合作大有帮助.

蜂巢可以被看作一个精简版的多细胞有机体. 这个有机体的每个体细胞携带相同的基因, 但体细胞不直接自我繁殖, 没有生育能力的工蜂同样如此. 通过生殖细胞卵子和精子, 有机体的体细胞相互合作完成其基因的传播.

将人类社会看作是目标单一的多细胞生物的观点是错误的. 人们大多会养育后代. 不过非亲缘关系的合作也广泛存在. 尽管自然选择使得我们群居 (在三百万年前我们的祖先就已如此), 但是我们各自的行动既不像肝细胞那样协调一致, 也不如社会性昆虫那么紧密无间. 人类的合作往往是基于个人利益所做出的选择.

我们的社会没有什么超级组织. 英国前首相撒切尔夫人曾一针见血地指出 “其实没有这样的社会”. 这可以作为 “方法论个人主义” (methodological individualism) 的战斗口号. 这个方法论通过分析参与个体的相互作用来解释自下而上的整体现象. 其所采用的数学工具是博弈论. 每个 “参与者” 都有各自的目标. 当然, 最终的结果可能完全不同于任何参与者的目标.

1.2　无 形 的 手

如果最终结果依赖于一些具有互不相同甚至相反利益个体的决策, 则可能会产

生激烈的冲突. 1651 年霍布斯 (Hobbes) 在其著作《利维坦》(*Leviathan*) 中声称, 利己的欲望将人类引入到 “人人为敌的战争之中”. 在缺乏抑制这些冲突的中央集权时, 人类社会充斥 “肮脏、贫穷、粗野、暴力和短视”. 同时代的法国作家帕斯卡 (Pascal) 与他同样悲观, “我们生而不平等, 因为每个人都为自己着想. 利己的欲望是战争、政治、经济等领域所有混乱的根源”. 利己被描述为万恶之源.

但一个世纪后, 亚当 · 斯密给出了另一种观点. 他认为一只无形的手在协调个体的利己行为: 人们在努力最大化自身利益的过程中, 也实现了整体利益的最大化. 利己的人在无形之中为公共利益做出了贡献. “在追求个人利益的过程中, 人们对社会的促进作用往往要比其刻意为之而更加有效.” 贪欲促成了客观上利他的行为. “我们能享用食物, 不应归因于屠夫、酿酒师以及面包师的恩惠, 而应归因于他们的利己动机. 我们应该诉诸他们的利己动机而非仁慈之心, 应该关注他们的获利而非我们的需求.”

早在亚当 · 斯密之前, 伏尔泰就在其著作《哲学通信》(*Lettres Philosohiques*) 中表达了类似的观点: “毫无疑问, 上帝完全可以把人类创造成只为他人着想的生物. 那样的话, 商人们将远涉印度去做慈善, 石匠们将无偿切割石块以取悦邻居. 然而上帝另辟蹊径 …… 正是由于相互的需要, 我们才对整个人类有贡献, 相互需要是一切商业的基础, 也是人与人永恒的联系.”

亚当 · 斯密 (Adam Smith)(十分了解伏尔泰) 非常清楚那只无形的手也有失效的时候. 因此, 他仅仅声称它往往会而不是必定会促进社会的整体利益. 今天我们知道, 在很多社会困境的问题中, 无形的手无法将个体自身的利益转化为整体的利益.

1.3 囚徒困境

要求两位参与者独自决定是否向对方捐赠. 如果捐赠, 则需支付五元, 对方将获益十五元. 显然, 如果参与者双方都选择合作, 即向对方捐赠, 那么他们将分别获益十元. 同样显然的是: 对每个参与者自身而言, 最有利的策略是拒绝捐赠, 即选择背叛. 无论对手选择合作还是背叛, 选择合作而舍弃五元都不符合你的利益. 如果对手选择合作, 你需要在获益十五元还是十元之间做出选择, 显然应该背叛; 如果对手背叛, 你需要在获益零元还是损失五元之间做出选择, 显然还是应该背叛. 下图描述了捐赠博弈的策略收益关系. 与此同时对方也处于相同的境地. 由于利己的欲望, 两位参与者都选择背叛而最终导致了不利于双方的糟糕结果. 无形之手怎么没有发挥作用呢? 如经济学家约瑟夫 · 施蒂格利茨 (Joseph Stilitz) 所说, “看不见是因为并不在此”.

		对手同意捐赠	对手拒绝捐赠
我的收益	我同意捐赠	10 元	−5 元
	我拒绝捐赠	15 元	0 元

这个奇怪的博弈是囚徒困境的一个例子. 该博弈是在两位参与者 (参与者 I 和参与者 II) 之间展开的, 每位参与者有两个选择: 合作 (cooperate, C) 或者背叛 (defect, D). 如果双方都选择合作, 他们将获得奖励 R, 这将优于双方都选择背叛所得到的惩罚 P. 但是如果一方背叛而另一方合作, 那么背叛方将获得收益 T (诱惑), 这甚至要高于双方都合作的奖励 R, 而合作方将获得最低的收益 S (傻瓜的代价). 于是有

$$T > R > P > S. \tag{1.1}$$

如前所述, 无论对手选择何种策略, 你都应该选择背叛. 如果参与者都旨在最大化自身的收益, 那么最终将得不到最优的结果. 这个结果是一个陷阱: 双方都不愿单方面将策略由背叛改为合作. 当然如果双方都选择合作, 那么将得到好的结果. 但是, 如果你一旦知道对方将选择合作, 获取更大利益的诱惑将驱使你采取背叛策略. 我们回到本节之初, 理性的策略方案只能是选择背叛, 这将陷入一个经济学的僵局.

		参与者 II 选择合作 (C)	参与者 II 选择背叛 (D)
参与者 I 的收益	参与者 I 选择合作 (C)	R	S
	参与者 I 选择背叛 (D)	T	P

这类博弈之所以被称为 "囚徒博弈", 是因为其在二十世纪五十年代初期首次被提出时是基于一个关于涉嫌共同犯罪的两囚徒故事. 为了使嫌疑犯坦白, 州检察官将两个囚徒隔离开并向他们分别做出承诺: 如果背叛同伙, 则可以获释 (作为污点证人), 而其同伙将面临十年的徒刑. 但是双方都明白他们不能同时成为污点证人, 如果双方都坦白, 那么他们都将服刑七年. 如果他们都保持沉默, 检察官会搁置审判, 将他们拘押一年. 这便是囚徒博弈的原型. 此处的诱惑 T 是成为污点证人, 奖励 R 则是经过审判 (最多被推迟一年) 后被释放, 惩罚 P 则是七年的徒刑, 而傻瓜的代价 S 是十年的监禁.

作为冷战时期的主要智囊团队兰德国公司雇佣的年轻数学家首先研究了这类博弈问题. 可能是两个超级大国面临的两难处境激发了他们的研究灵感. 苏联和美国如果联合核裁军将会使双方都受益. 但诱惑是各方都想保留一些原子弹而让对方销毁核武器库, 这将引发昂贵的军备竞赛, 从而导致糟糕透顶的结果.

1.4 雪堆博弈

反映利己弊端的社会困境不只囚徒困境一例. 另一个例子是所谓的雪堆博弈.

假设实验组织者向两位参与者承诺, 如果他们愿意支付总共三十元的费用, 那么将各获得四十元的收益. 两位参与者独立决策是否愿意支付费用, 并且双方明白, 如果他们都愿意支付, 那么将对半分摊总费用. 显然的策略方案是: 两位参与者分别投资十五元, 然后各自收回四十元, 即各自净获益二十五元. 但假如某一位参与者拒绝支付费用. 此时, 对另一位参与者的建议是独自承担三十元的费用. 这样的话, 他还是可以获得十元的收益. 然而该决策难以让人接受, 因为此时对手将不支付任何费用而获得四十元. 这个博弈中的可选决策可以描述成合作 C (意味着愿意承担费用) 和背叛 D(意味着拒绝承担费用). 如果参照公式 (1.1) 的记法, 将对应收益值记为 R, S, T 和 P, 则有

$$T > R > S > P. \tag{1.2}$$

由于顺序的微小变化 (仅仅 S 和 P 交换了顺序), 不管对手策略如何而总是背叛, 已不再是最佳选择. 如果对手选择背叛, 那么你最好选择合作. 实际上, 对每个参与者而言, 最佳选择是采取与对手相反的策略. 此外, 双方都知道若能成为背叛者将会更好. 如果双方都坚持自己的最佳选择, 那么最终会陷入糟糕的结果. 于是有一方会作出让步, 至此达成妥协, 博弈也随之结束.

		对手同意投资	对手拒绝投资
我的收益	我同意投资	25 元	10 元
	我拒绝投资	40 元	0 元

雪堆博弈源于这样一个故事: 两个司机被大雪堵住了道路, 无法回家. 如果想回家, 他们必须铲除路上的积雪. 公平的方案是双方都去铲雪 (假设他们的后备箱中都有一把铲子). 若假设某一位司机固执地拒绝铲雪, 另一位司机也可拒绝铲雪, 这就意味着他俩要在野外度过寒冷的夜晚. 此时, 尽管对方不劳而获, 另一方还是应该选择去铲雪.

1.5 重复困境

囚徒、超级大国以及经济实验中的测试对象看似与我们的日常生活相去甚远, 但在日常生活中, 大多数人每天都会经历许多类似的轻量级经济往来. 甚至远在市场和货币出现以前, 人们就在家庭内部, 团体内部以及邻里之间忙于频繁的给予和获得的互动, 并时刻面临给少取多的诱惑.

捐赠博弈中的不切实际之处, 并不在于它的收益结构, 而在于其对两位参与者仅博弈一次就分道扬镳的假设. 日常生活中我们大多数的交互活动是在家庭成员、同事以及其他经常见面的人之间展开的.

到目前为止, 我们所讨论的是所谓的单回合博弈. 现在假设两位参与者重复博

弈多个回合. 显然参与者如果禁不住诱惑而背叛对方, 其必定会预料对方将报复. 一方参与者在当前回合所采取的策略将影响对方在下一回合的策略选择.

假设参与者重复六次捐赠博弈, 这将会提高合作的可能吗? 根据逆向归纳法原理并不能肯定. 可以先只考虑第六回合, 即最后一回合. 该回合之后不再有后续回合, 而此前的回合也已经尘埃落定. 因此, 该回合可以被视作一个孤立的回合, 于是它就退化成单回合博弈. 正如我们所见, 从利己的角度出发, 双方应互相背叛. 这就是所谓的最后一回合效应. 博弈双方都明白他们无法改变这一结局. 因此, 他们干脆接受这一事实, 在决策时忽略最后一回合, 只考虑此前的五个回合. 于是按前面的讨论同样可以得出第五回合也应该相互背叛. 以此类推, 根据逆向归纳法原理最终得出双方始终不应该选择合作. 参与者虽面对每一轮可以吐出十元钱币的机器, 但利己欲望导致他们错失良机. 这是相当匪夷所思的. 显然, 一个明智的做法是在首回合中采取合作以做出姿态, 向对方表明你不愿采用冷酷的逆向归纳法逻辑来行事.

前面的问题是无关主旨的. 实际上, 人们在日常活动中事先很难知道哪一轮是最后的回合. 通常后续回合可能继续发生 —— 此所谓未来的阴影. 例如, 假设实验人员告诉参与者, 每一回合结束后将掷一次骰子. 如果点数为六, 则博弈终止, 否则博弈继续. 这样博弈的回合数是随机的. 可能下一回合就结束, 也可能再持续二十回合. 尽管博弈的平均回合数是六, 但你只能选择合作, 否则不免要担心遭到报复.

与单回合囚徒困境相比, 在重复囚徒困境中不存在一种策略在各种情况下都是最优的. 如果你的对手采取无条件的背叛或者合作, 则显然背叛是最佳选择. 然而如果是一个难以对付的对手, 他只有在你一直选择合作时才和你一样总是合作, 而一旦遭到你的背叛, 他将在此后无情地永远选择背叛, 此时你最好在每一回合都采取合作. 实际上, 如果你选择背叛, 你占了这类对手的便宜而多获益五元, 但你将在未来回合中失去所有的奖励, 再也不可能得到正的收益. 由于博弈后续回合数的期望是五, 这就相当于你放弃了五十元.

重复的雪堆博弈又会怎样呢? 容易发现, 如果在所有回合中, 两位参与者选择共同承担费用或者交替独自承担费用, 最终他们各自的平均收益是相同的. 但他们可能达成这样对称的方案吗? 经过几轮回合后, 真的会有一方作出让步, 而甘愿成为被对方利用的合作者吗? 在此情形下, 他们的总收益与双方都合作时是相同的, 但各自的收益却高度的不均衡.

1.6 排 名 赛

对于重复博弈, 当你知道没有最优策略时, 你会采取什么策略呢? 大约三十年前, 政治学家罗伯特 · 阿克塞尔罗德 (Robert Axelrod) 通过发起一场计算机排名

赛研究了这个问题. 人们可以提交相关策略的程序. 这些程序以循环式赛制进行两两博弈: 每个策略与其他策略随机相遇, 博弈 200 回合 (参与者事先不知道博弈回合数, 因此逆向归纳法在这里无用武之地). 有些策略方案设计相当复杂, 它们探测对手的响应并试图利用其弱点. 但大获全胜的赢家却是所有策略中最简单的一个, 它是所有报复策略的典型代表, 即称为以牙还牙 (Tit for Tat, TFT) 的策略. 采用 TFT 策略的参与者在首回合选择合作, 此后它只是简单地重复对手在前一回合的选择.

TFT 策略夺冠着实让人大吃一惊. 这看起来几乎是矛盾的, 因为在重复囚徒博弈中, TFT 策略的收益总是不高于其对手的收益. 实际上, 在系列回合中, TFT 策略从没机会领先对手. 只要双方都选择合作, 那么他们的收益相等, 如果对方首先选择背叛, 则对方将获得 T, 而 TFT 策略仅获得 S, 但在随后的回合中 TFT 策略将不再失手. 只要对方继续采取背叛, 双方后续回合的获益总是相同, 即 P. 一旦对方恢复合作, TFT 策略就会追回收益差距, 同时双方合作的局面随之恢复. 在博弈的任一阶段, TFT 策略的累积收益要么与其对手持平, 要么只落后 $T-S$. 但在阿克塞尔罗德排名赛中, 每个策略的总分是其与其他策略两两博弈的收益结果的累加, 据此排名, TFT 策略胜出.

阿克塞尔罗德发现十六个策略中的八个策略是善良的, 即它们不首先背叛. 而正是这八个策略占据了前八名. 善良的人赢得了胜利! 实际上, 阿克塞尔罗德发现, 如果另一个策略参赛, 它应该夺冠. 这个方案比 TFT 更为善良, 它就是两报还一报 (Tit for Two Tats, TFTT). 该策略规定仅仅在对手连续两次背叛后才选择背叛. 当阿克塞尔罗德第二次召集计算机联排名赛时, 有六十四套策略程序参加了比赛, 其中有一个正是 TFTT 策略. 但这个本被视作第一场排名赛夺冠热门的策略, 出人意料的在第二次比赛中只得到了第二十一名. 第二次比赛的冠军还是被简单的 TFT 策略摘得. TFT 不仅仅善良, 同时它鲜明, 易怒, 宽容和稳定. 正是这些特质促成了它最终胜出.

1.7 虚拟社会

阿克塞尔罗德计算机排名赛取得的成功引发了计算机模拟的热潮. 基于个体的虚拟社会模型极大地扩展了博弈论的研究范围. 虚拟社会由虚构的个体构成, 每个个体由一个程序为其设定了策略行为. 这些个体随机相遇而相互进行重复囚徒博弈, 然后继续与其他个体相遇博弈. 最终比较各个策略博弈收益的累计结果. 通常这类排名赛会更新虚拟群体数量. 这意味着个体会产生后代, 也就是会有虚拟的个体继承他们的策略行为. 那些拥有高收益的策略所对应的个体将产生较多的后代, 于是成功的策略得以传播. 另外还有一种机制, 它不是基于策略行为的继承, 而是

允许新的个体通过复制去学习那些成功个体的策略行为. 在这类基于个体的模拟实验中, 各种策略对应个体的频度将随时间不断变化. 也可以不时地引入少量新策略的个体, 观察他们是否能够侵入成功. 在第 2 章中, 我们将通过描述相关的数学背景来研究这类模型.

假设一个群体里面只有两种策略 TFT 和 AllD(always defects). 如果两个 TFT 策略相遇, 则其平均收益为六十元 (对应六个回合的合作). 如果 TFT 策略与 AllD 策略相遇, 后者获益十五元 (通过在首回合背叛对方), 前者则损失五元. 如果博弈双方都采用 AllD 策略, 则他们一无所获.

		对手采取 TFT 策略	对手采取 AllD 策略
我的	我采取 TFT 策略	60 元	−5 元
收益	我采取 AllD 策略	15 元	0 元

参与者最好是根据对手采取的策略在两个策略中做出选择, 即采用 TFT 策略来应对 TFT 策略的对手, 而对 AllD 策略的对手则应该采取 AllD 策略. 但在基于个体的模型中, 虚拟参与者别无选择, 他们只能固守其策略, 并且事先也不知道对手的策略. 显然, 期望收益依赖于虚拟群体的组成结构. 如果大多数参与者采取 TFT 策略, 则 TFT 策略是较好的选择; 而如果群体里充满了背叛者, 则 AllD 策略是较好的选择. 在后一种情况, 参与者陷入了一个社会困境. 这类结构的博弈也被称为猎鹿博弈 (Staghunt games). 虚拟群体的进化将使得所有个体采取相同策略. 具体的最终进化结果依赖于初始条件. 显然, 如果 TFT 策略的数量超过 10%, 那么他们将最终胜利. 如果下一回合发生的概率接近 1, 即后续回合数的期望值很大, 那么即使是少量的回报者也能最终战胜背叛者.

计算机模拟表明, 一个纯粹的 TFT 世界并非历史的终结. 实际上, AllC(always cooperates) 策略的参与者就可以侵入, 因为此时侵入者与抵抗者收益相同. 即, 如果将少量的 AllC 策略引入到原来全部都是 TFT 策略的环境里, 他们与占绝大多数的 TFT 策略收益相同. 实际上, 在某种情况下, AllC 策略甚至具有一定的优势. 无条件策略应该比有条件策略更容易实施. 更重要的是, 两个 TFT 策略相互博弈时, 如果出现误操作或者误解, 那么将因为策略本身以眼还眼的冷酷性而使双方陷入一个成本昂贵的交替背叛的循环之中. 在计算机模拟过程中, 这类错误可以被杜绝, 但是在现实生活中则无法完全避免. 由于自身的误操作或者对方的误解而造成的错误非常普遍. AllC 策略对错误不敏感, 即他与 TFT 策略者或其同类在博弈过程中发生的错误, 可以立刻在下一回合中得到纠正.

如果基于个体的模拟像现实生活那样偶尔出错, 那么纯粹的 TFT 世界不会长久维持, 诸如 AllC 等宽容的策略将得到传播. 但是一旦出现了相当数量的 AllC 策略, AllD 策略将获得胜利. 模拟群体的重复囚徒博弈进化过程观看起来令人如痴

如醉. 实验结果以一种令人惊奇的方式依赖于基于个体模拟中反复试错的策略的范围. 一个经常夺冠的策略是巴浦洛夫 (Pavlov) 策略, 该策略总是首先合作, 此后只有当前一回合博弈双方采取相同策略时, 才选择合作. 在第 3 章中, 我们将分析与基于个体的模拟相关的某些博弈理论.

1.8　回报的冠军

计算机模拟排名赛引发了回报问题研究的高潮. 但是它在多大程度上与现实相关, 还是仅仅与思维层面的实验相关? 如果 TFT 策略是那么的完美, 那么它应该得到广泛传播才对. 进化生物学家以及动物行为学专业的学生发现了少量的候选策略行为, 但没有一个是被普遍接受的. 在现实世界中, 很难保证收益值 (在动物世界里, 收益值表现为成功繁殖的能力) 遵循不等式 (1.1) 中参数的大小顺序. 通过几个回合的观察, 我们甚至很难判定具体采用了哪个策略. TFT 只是无数可能的策略之一. 此外, 许多现实生活中的合作提供了大量其他解释的空间, 如亲缘选择.

经过几十年的研究, 最终的结果发人深省. 除了灵长类动物以外, 以牙还牙策略的确凿例子十分罕见. 另一方面, 大量的证据表明人类是回报的冠军. 不仅大量的心理测试和经济实验显著印证了这个观点. 脑部成像似乎也支持这个观点, 它显示我们的大脑皮质的某个部分专门负责对收支决策所需的计算做出处理, 以及对感知到的不平衡做出情感反应. 另外, 人类在移情方面具有杰出的天赋——即换位思考的能力. 我们不仅有模仿他人行为的天性, 同时我们还有洞悉他人动机的能力.

就人类本性而言, 报复现象随处可见. 从小孩子回踢弄伤自己的无生命的物体可见人类的报复心理是如此强烈. 更重要的是我们会与互动的陌生人甚至旁观者 (即所谓的第三方) 进行换位思考, 这开启了间接回报的研究.

1.9　引入第三方

你可能听过这样的故事, 一位年长的教授一丝不苟地参加同事的葬礼, 其原因是 “如果我不参加他们的, 其他人也不会来参加我的”. 显然, 直接回报在此是不起作用的. 可以预见, 应该会有很多人来参加这位教授的葬礼. 他之前坚持参加同事的葬礼所表现出的尊敬, 将从第三方而不是那些已故同事那里得到回报. 这就是所谓的间接回报.

对于直接回报, 帮助的举动将从受助方获得回报. “我帮你挠痒因为你帮我挠痒.” 但在间接回报中, 帮助的举动将从第三方而不是从受助方获得回报. “我帮你挠痒是因为你帮别人挠了痒.” 这看似很难理解. 但从原理上说, 这可以使合作得以持续, 或者更准确地说, 可以防止合作被利用而导致毁灭.

实际上, 剥削者将无偿接受帮助而不给予任何回报. 如果所有人都这样, 合作将不复存在. 因此, 必须抑制这样的剥削行为. 一个显而易见的方法是不帮助这些剥削者, 而只帮助合作者. 感觉上这种规则还是有矛盾的地方: 如果你拒绝帮助一个剥削者, 第三方可能将你也当作剥削者, 这样你将会被他们排斥在合作之外. 但是, 我们将在第 4 章看到这类间接回报的机制可以持续下去.

如果第三方能够将剥削者的背叛行为与惩罚剥削者而采取的报复行为予以区分, 那么这种拒绝帮助剥削者的行为则无任何风险. 这要求旁观者即便未直接参与博弈, 也必须能够判定行为正义与否, 即好或坏.

仔细研究发现存在着许多种可能的评价准则, 有些准则要优于其他准则. 所有的评价准则都需要其他参与者的大量信息. 处理这些信息的能力, 可能在直接回报的过程中已经得到进化. 在你开始一系列重复博弈之前, 事先了解可能的对手在其历史博弈中的策略选择将是非常有帮助的. 从这个角度来看, 有未来潜在的博弈伙伴见证的直接回报将产生间接回报, 进化生物学家理查德 · 亚历山大如是说. 然而直接回报有赖于重复, 而间接回报有赖于声誉. 对于前者, 我们需要认识对手, 而对于后者, 则需要知道对手. 对于直接回报, 我们需要知道的是一张脸, 而对于间接回报, 我们需要知道的是一个名字 (大卫 · 黑格, David Haig).

易买 (Ebay) 要求买家在每次交易完成后对卖家的服务做出评价. 易买会员在最近十二个月的累计评价等级是众所周知的. 这些比较原始自然的评价准则似乎足以完成声誉的建立. 人类历史上存在大量公开评价的例子: 割掉大拇指以标识一个小偷, 剃光头以显示一个不贞的女人, 颁发勋章以宣示一名英雄. 在远距离交易出现后声誉机制依然起着十分重要的作用.

如果是在小型社区里, 自己直接参与博弈以及旁观他人博弈就足以保持间接回报. 但在大型社区里, 个体通常需要依赖于第三方的信息. 闲聊总是信息传播的重要工具. 很可能我们的语言本能和道德观念在共同进化.

1.10 道德情操与道德危机

亚当 · 斯密非常了解日常经济决策过程中道德判断的重要性. 他在写《国富论》之前便完成了《道德情操论》这部著作. 其后的经济学家试图忽略道德情感的因素. 但是, 今天人们已经广泛接受一个观点, 那就是信任是各种交易的首要条件. 信任被誉为 "社会生活的润滑剂". 不同社区具有不同的信任水准. 经济活动中严格的道德基础和双方认可的法律规则是表征一个国家财富的主要指标, 其重要性超过人口数量以及矿产资源.

人类偏爱信任的天性在所谓信任博弈的实验中展露无遗. 信任博弈建立在捐赠博弈的基础之上: 第一步, 捐赠者 (或投资者) 从实验组织者获得一份捐款, 并且

决定是否将其中一部分给予接受者 (或托管者), 注意其给予部分将为托管者带来三倍的利润: 即投资者每 1 元的投资将为托管者带来 3 元的收益; 第二步, 托管者可以向投资者以一比一的方式返还部分利润, 即托管者 1 元的支付将为投资者带来 1 元的收益. 博弈由此结束, 并且参与者清楚他们将不再相遇. 显然一个利己的托管者将不会返还任何收益给投资者. 而一个利己的投资者也不会向托管者进行任何投资. 然而在实际的实验中, 参与者往往会选择投资和返还, 从而取得共赢的结果. 信任博弈将在第 5 章中分析, 结果毫无意外地表明对声誉的关注将起到重要的作用.

现实中的许多经济活动包含了信任博弈的基本元素, 假如我把资金委托给一名基金经理, 我会期望获得一份正利润, 而基金经理也期望获得一份收益. 对一个基金项目而言, 良好的声誉是极其重要的无形资产. 不能返还资金的银行家将面临双重麻烦: 一方面老客户会坚持要求取回资金, 另一方面不会再有新客户愿意把自己的血汗钱委托给他. 此时, 直接回报与间接回报同时发生作用.

经济学家和社会学家越发关注间接回报, 这是因为在全球化的今天, 相距遥远的买卖双方之间的一次性交易日益频繁. 它们将取代在亲戚邻里之间的传统交易. 我们生活中相当一部分时间用于和许多陌生人打交道, 并且许多交易不再是面对面进行. 电子拍卖以及其他形式电子商务的业务增长在很大程度上依赖于声誉和信任. 而信用被利用的风险引发了经济学家所谓的道德危机. 声誉的有效性如何, 特别是在信息不完整的时候?

另一方面, 进化生物学家对人类社会的出现过程感兴趣. 人类之间相当部分的合作是基于道德情感的, 比如, 对小偷的愤怒以及众所周知的在帮助他人之后所感到的精神愉悦. 有趣的是, 人们不仅对自身直接参与的活动也对旁观到的第三方之间的行为感受强烈. 他们根据道德规范行事. 这些规范在很大程度上因文化的差异而不同. 但是道德规范的能力是人类所共有的, 而对其他物种几乎无迹可寻.

显而易见, 一个有机体是按照对自身繁殖有利与否而将事物划分为好与坏. 我们变得喜欢高热量食物以及性生活是因为这提高了我们生存和繁殖的适合度. 相反地, 厌恶、饥饿以及疼痛作为警示信号提醒我们远离危险. 根据自身的感受是很难准确判定第三方之间的行为是 “好” 还是 “坏”. 但是, 用于表达自身快乐或不适的术语 “好” 与 “坏” 也被用于评价第三方之间的行为. 这种语言上的癖好展现了人类无比惊人的移情能力, 同时也反映了高度发达的认知及抽象能力.

1.11　最后通牒博弈

一系列的经济实验表明间接回报是有作用的. 参与者彼此相互了解越多, 就越有可能相互帮助. 似乎存在明显证据表明参与者都关注自己的声誉. 但有趣的是, 当

没有人旁观并且所作所为不会影响声誉时, 很多参与者还是会倾向于提供帮助, 虽然程度上会稍微低一点. 此外, 他们如果以前接受过帮助, 则更有可能提供帮助. 从利己观点出发很难解释这一点. 它可能是某种不适应所产生的结果. 如果有人为你扶住门, 那么你很有可能会为下一个人扶住门, 这往往是出于一种模糊的感激心态. 很可能就是这些类似的间接回报机制在其他社会和经济环境中发挥作用.

揭示人们同情他人的一个典型实验是最后通牒博弈. 在这个博弈中, 两个匿名参与者被随机指定为提议者和接受者. 实验组织者给予提议者十元, 要求他提出分配方案, 并让接受者选择. 即, 如果接受者同意分配方案, 那么就按照这个提议方案进行分配, 游戏就此结束. 但是如果接受者拒绝这个方案, 则游戏也结束. 实验组织者收回十元, 这两名参与者将一无所获. 这个博弈的特点就是: 没有讨价还价, 没有第二轮.

很明显, 接受者应该接受任何正的分配方案, 因为这总比没有好. 因此, 一个利己的提议者应该提供尽可能少的份额给对方. 然而, 在实际的实验中, 大多数参与者都给出了公平的分配方案: 给对手的份额在百分之四十到百分之五十之间. 低于百分之二十的分配方案偶尔被提出, 但接受者往往予以拒绝, 提议者似乎也预料到了这一点.

在大多数情况下, 拒绝伴随着愤怒的情绪. 脑部成像显示不公平的提议会引起大脑两个区域的活动: 一是位于大脑的左额叶, 它通常与理性决策有关, 而另一区域则位于更深层的纹状体, 它与情绪反应相关. 这两个区域之间的拉锯战反映了两种决策在大脑里的激烈斗争, 即 (a) 接受低份额分配方案, 理由是它总比没有的好; (b) 告诉不公平的提议者 “见鬼去吧”. 甚至在响应者未明确做出决策前, 这两个区域的大脑活动强度已经预示了决策的结果.

人们利用最后通牒博弈反复实验, 同时研究了大量的变种版本. 例如, 由计算机来扮演提议者, 则接受者会毫无怨言地接受较小的报价. 如果通过一种技能游戏的比赛 (而不是抛硬币) 来决定谁扮演提议者, 那么较小的报价也可能被接受, 因为这仿佛是提议者赢得了保留较大份额的权利. 此外, 如果有多名接受者相互竞争, 那么提议者也知道小的报价极有可能被接受.

1.12 公平准则

大量的研究利用最后通牒博弈来考虑许多小型团体内的公平准则, 这些小型团体包括狩猎采集者, 游牧民族, 刀耕火种的农民. 平均报价因文化差异而不同. 值得注意的是, 来自大城市的报价最为公平, 大自然之子 (Mother Nature's Son) 并不比城市中的骗子甚至经济学的本科生高尚多少. 但是平均报价总是偏离最小理论值. 公平准则似乎广泛存在, 并可能普适. 但它是如何产生的呢?

声誉再次成为一种可能的解释. 一旦人们知道你会拒绝不公平的报价, 那么他们在向你提出报价之前会慎重考虑. 长远来看, 拒绝接受不公正报价所带来的好处将远大于你的损失, 相比而言, 因拒绝而失去的份额是那么地微乎其微. 在第 3 章中, 一个简单的数学模型将揭示人们对声誉的关注是如何促成公平准则的建立. 矛盾的是, 这只在如下情形才有效, 即提议者通常愿意提出公平的报价, 只是偶尔在可以逃避惩罚时经不住诱惑而提出不公平的报价. 公平准则并非人们所固有, 且可能被机会主义的私欲所颠覆, 因此它必须在群体中得到捍卫.

关于这一点在真实的实验 (相对于基于个体的计算机模拟) 中如何呢? 很容易设置两个不同的规则, 在大数量人群中随机配对进行匿名的最后通牒博弈. 一种规则, 参与者进行十回合博弈 (当然是与不同的对手), 大家都不知道对方在前面回合中的策略选择. 另一种规则, 参与者在博弈中的选择向所有人公开. 显然只有在第二种规则下才有可能建立起拒绝低报价的声誉. 结果非常清楚: 不公平的报价显著减少. 提议者似乎预料到这一点, 即接受者会担心自己接受低报价的软弱行为被大家知道而遭到利用.

如果最后通牒博弈中的接受者拒绝不公正的报价, 他们会尽力让所有人知道这一点. 在自然环境里, 情绪化的反应往往会吸引人们的注意力. 愤怒总是大声的.

虽然如此, 然而参与者即使知道博弈结果将被保密, 最后通牒博弈中的报价仍然会比较公平. 这似乎令人费解, 很可能是因为潜意识里参与者很难相信真的不会为人所知. 在进化的历史上, 祖先们始终生活在小型社区里, 想瞒着部落成员或是村庄居民是极其困难的. 此外, 绝大多数人相信有一个人格化的神灵在时时刻刻盯着他们, 这种被监视的感觉根深蒂固且被广泛接受.

心理学家设计了巧妙的实验, 证明了这种被监视的感觉很容易被唤醒. 例如, 坐在小隔间里的参与者面对着一台电脑, 哪怕屏幕上只是显示一只眼睛, 他也会深受影响. 尽管他们清楚这只眼睛完全是象征性的, 但还是会对它做出反应. 还有一个简单且精彩的实验: 在一个英国大学系部的自助餐厅里, 只是在 “诚实箱” 旁边的墙上张贴了印有一双眼睛的照片, 就足以将支付咖啡和饼干的教职员工人数提高百分之两百以上. 显然, 很容易触发对被监视的关注. 值得强调的是, 对我们这个物种而言, 眼睛具有独特的启示作用: 由于虹膜周围的颜色是白色, 眼睛注视的方向特别引人注目. 此外, 不管是一人还是多人围观, 测试对象都会做出同样的反应. 这说明他们至少在潜意识里认为消息将会通过闲聊传播出去. 有一个目击者就足够了.

1.13　公共利益博弈

到目前为止, 我们讨论的诸如囚徒困境、雪堆博弈、信任博弈、最后通牒博弈

都是二人博弈. 然而众多的经济活动包含了大量的参与者. 此时, 回报的提法就显得有些问题. 如果群体里既有合作者, 又有背叛者, 那么你会对谁进行回报呢? 这使得社会困境问题更加复杂纠结.

所谓的公共利益博弈提供了这类困境的例子. 下面是一个典型的实验案例: 事先分配给六位匿名参与者每人十元, 并让他们向公共资金池中投资一部分金额. 他们都知道, 实验组织者会将公共资金池中的总投资额转化为三倍的收益, 并且将这份 "公共利益" 在六位参与者之间平分.

显然, 如果每个参与者都全额 (十元) 投资, 那么对所有人都有好处, 每人将获得三十元. 如果仅有一位参与者拒绝投资, 而其他参与者全额投资, 则每人都将获得二十五元的回报, 而那位未作任何投资背叛者还保留了初始分发的十元, 最终持有三十五元的总收益, 高于其他参与者十元.

对具体的参与者而言, 每投资一元只能带来五角的回报. 因此对一个追求利益最大化的利己参与者而言, 他应该拒绝做任何投资. 但如果每位参与者都拒绝投资, 那么他们将错失增加收益的大好机会.

在实际的实验中, 多数参与者的平均投资额为初始分发金额的一半, 甚至更多. 个体的投资额具有很大差异, 但是许多参与者看起来都是谨慎下注的. 然而, 随着博弈持续几个回合, 投资额度将越来越低, 最终可能降为零. 机制是明确的, 如果有人注意到自己投入的比别人多, 就会觉得自己被剥削而随之降低投入. 这又会导致其他参与者感到被剥削而随之降低投入, 于是, 合作趋于崩溃.

在重复囚徒困境博弈中, 类似 TFT 的策略可以用来报复背叛者. 但是这个回报策略在重复公共利益博弈中是失效的. 实际上, 当你拒绝投资时, 你的朋友 (合作者) 和敌人 (背叛者) 都受到打击: 你的报复不只是针对那些背叛者, 而是会影响到公共利益博弈的所有参与者.

在经济活动中, 基于共同努力, 联合投资的经济活动广泛存在. 其潜在的社会困境经常被描述为多人囚徒困境, 搭便车 (free-rider) 或者公共用地的悲剧 (tragedy of commons). 公共用地可以是一块村民们都可以使用的牧场. 悲剧则是源于这个牧场经常会被过度放牧而遭到毁坏. 今天公共用地所剩无几, 但悲剧依然伴随着我们: 海洋便是我们新的公共用地. 从小范围来看, 多数的公共厨房也会发生这样的悲剧, 联合公司以及公共资源为小偷和背叛者提供了诱人的机会.

1.14 惩罚或毁灭

如果你无偿使用公共交通资源, 或者逃税漏税以及违章停车, 那么你有可能被抓住并罚款. 各种司法和仲裁机构以及道德压力在努力促使我们保持较高的合作程度. 于是搭顺风车问题有一个简单的解决方案: 可以通过激励机制来增强合作, 即

对个人的合作或背叛行为给予相应的奖励或惩罚. 然而法制环境依赖于高度发达的社会.

事实表明在没有这样的环境时, 个人往往愿意主动去扮演惩罚者的角色. 以下的实验巧妙地展现了这一点. 每回合公共利益博弈后, 每个参与者被告知对手的投资额度, 他可以据此实施惩罚. 如果被惩罚, 则被罚者的账户将被扣除三元, 该罚金由实验组织者收回, 而不是补偿给惩罚者. 另一方面, 这些惩罚者需要为惩罚的实施支付一元的费用. 这个费用可以理解为惩罚他人往往是一个有代价的行为: 需要时间和精力, 而且经常还要冒一定的风险.

在经济实验中, 尽管自身会产生一定的成本, 参与者往往还是愿意去惩罚. 很多参与者也期待有人去实施惩罚. 由于受到惩罚的威胁, 有惩罚机制后平均的投资度要高于之前没有惩罚的时候. 更为突出的是, 如果博弈持续好几个回合 (每个回合后包含公共利益博弈以及随后的惩罚环节), 那么投资度会越来越高以至于达到一个惊人的高度, 如图 1.1 所示, 惩罚显著地提高了合作的程度.

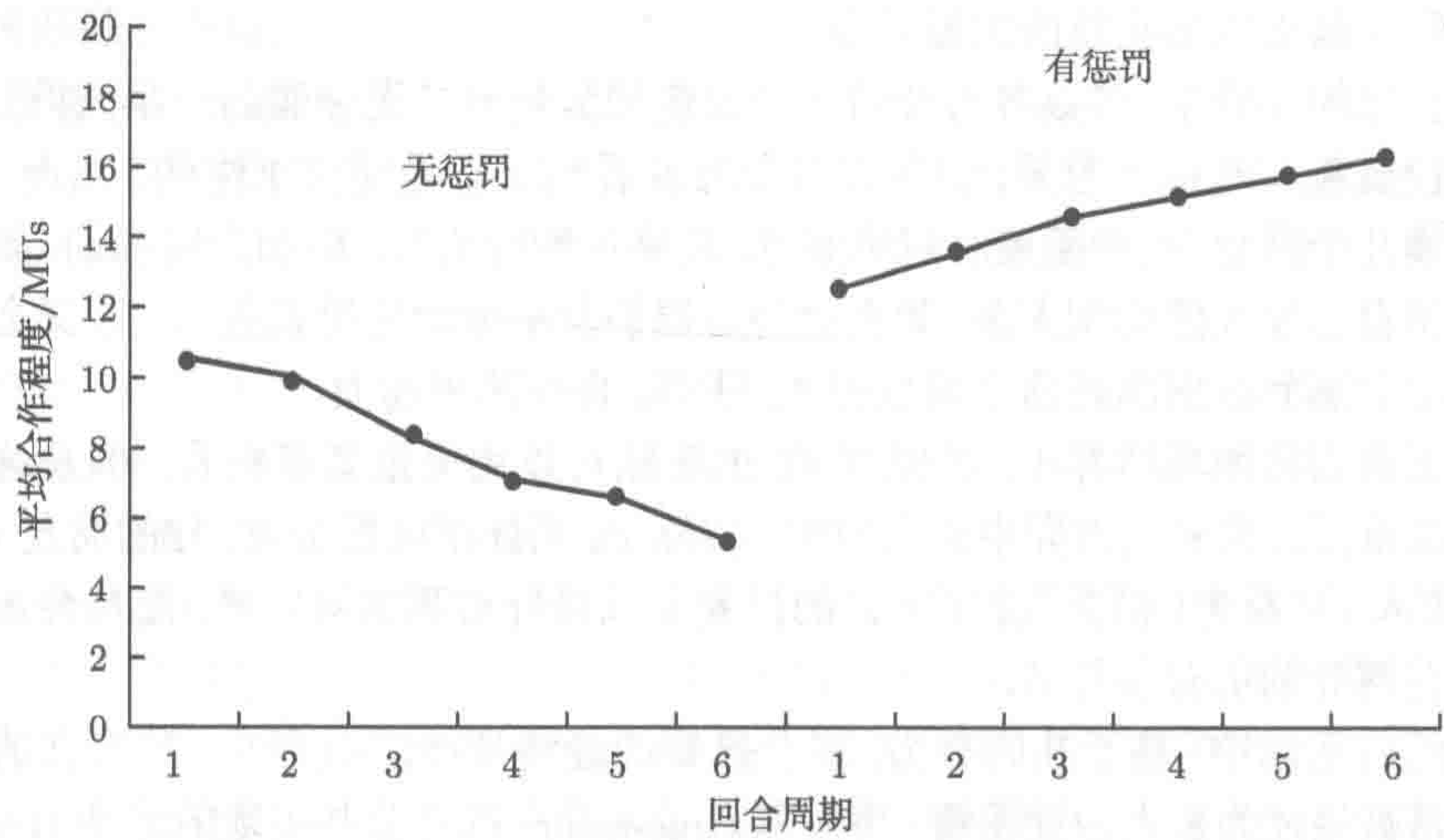


图 1.1　公共利益与惩罚. 在图示 12 回合的每一回合中, 博弈的四位参与者都是从 240 人中重新选出. 参与者在每回合之初得到二十货币单位 (money units, MU), 并决定投资额度. 他们知道其投资将翻番并在四位参与者之间平分. 在带有惩罚的规则下, 参与者可以惩罚对手, 罚金将由实验组织者收回. 三个货币单位的罚款需要一个货币的费用. 参与者知道他们与对手只会相遇一次. 图示为每回合的平均投资额度 (Fehr, Gächter, 2002)

既然有成本, 为什么人们还会去惩罚呢? 第一个解释是显然的: 惩罚者希望通过惩罚背叛者来改变他们, 并由此期望在随后的回合中, 这些被严惩的参与者会提高投资度, 从而可以补偿自身因惩罚背叛者而造成的费用损失. 但这看起来只是答案的一部分. 实际上, 在一个有大量参与者的博弈中, 每一个回合中的六位参与者

都是经过重组而成, 并且他们知道相互之间不会再次相遇. 通过实施惩罚, 他们可能会将背叛者改变为合作者. 但是惩罚者知道, 这些受罚而改变的背叛者在未来提升的投资度只会让其他人受益, 而不是自己. 这种惩罚明显是一个利他的行为.

这是一个惊人的结果. 在没有惩罚机制时, 公共利益, 即三倍的投资回报不能实现; 而有惩罚机制时, 出于利己考虑, 参与者应该放弃昂贵的惩罚. 在缺少司法环境时, 有些参与者愿意将"法律掌握在自己的手中", 这被称为"同级惩罚"(peer punishing).

是什么促使这些参与者去惩罚背叛者? 大概我们只需关注普遍认同的间接回报机制. 如果参与者因为自己觉得受到剥削而实施惩罚, 这就是直接回报在发挥作用; 而如果因为看到别人受到剥削而实施惩罚, 则是间接回报在发挥作用: 人们往往愿意作为第三方来进行报复.

1.15 二阶搭便车

如果参与者群体中有一部分人愿意去惩罚剥削者, 那么背叛显然是没有意义的. 拒绝投资而讨得的便宜不能抵消预期的罚款, 背叛者将完全承受大多数人对他的惩罚. 如果惩罚者 (自己合作, 并且对不合作者实施惩罚) 在一个群体中确立了地位, 那么他们可以抵制背叛并捍卫合作.

但由惩罚者构成的群体会被那些只合作而不实施惩罚的参与者破坏. 这类新来者与原有的惩罚者有相同的收益, 因而他们可能在随机波动下得到缓慢的扩散. 实际上, 如果偶尔有一些背叛者侵入, 当然他们立即会遭到惩罚者抵制. 此时, 新来者因为节省了惩罚成本而得到了比惩罚者更好的收益. 这类参与者被称为二阶剥削者, 他们搭了惩罚者所实施惩罚的顺风车. 这样, 此类参与者将得到扩散, 这意味着最终抵制背叛者的惩罚者将越来越少.

惩罚可以被看做是对社区的一个服务, 即公共利益. 长期来看, 这类二阶剥削者扰乱了公共利益博弈中贡献的实施. 最终, 这两类合作者, 即惩罚者以及二阶剥削者最终都将被背叛者淘汰.

可以想到的补救方法就是二阶惩罚: 它惩罚的对象不仅包含一阶剥削者, 即那些背叛者, 而且还包含二阶剥削者, 即只合作但不实施惩罚的参与者. 然而这又会导致三阶剥削者. 如果足够高级的惩罚者在主导整个群体, 那么其中的背叛者会很少. 在严厉的高级惩罚者面前, 那些低级的惩罚者鲜有机会暴露出自身的局限性. 于是他们的数量会在随机波动下增加, 从而侵蚀整个系统.

看来声誉再次拯救了这个危机. 如果知道参与者群体中存在惩罚者, 那么人们可能会抵制住背叛的诱惑. 人们在表露对剥削者的情绪反应上存在一种优势. 这些内容将在第 5 章中讨论. 这与最后通牒博弈类似. 实际上, 接受者拒绝一个不公平

的报价时, 实际上是对提议者的一个有效惩罚, 报价越公平, 惩罚者的惩罚实施成本越低, 而对背叛者的惩罚力度则越大.

另一种类似的机制是正向激励, 如果每一轮公共利益博弈结束后, 大家可以拿自己的一部分收益对贡献度高的参与者予以奖励, 则可以促进合作程度的提高. 同样, 这个机制会受到那些只合作不奖励 (即从别人的奖励行为中受益) 的参与者的威胁. 这类参与者显然搭了那些自己掏钱的奖励者的顺风车, 并最终破坏基于激励机制的系统. 如果奖励者可以赢得声誉, 则他们的团体将获得更高水平的合作.

在有一点上, 奖惩之间存在差别. 如果每个人都积极合作, 那么惩罚者无事可做, 但奖励者却仍需慷慨解囊. 在这一点上惩罚显得更为有效, 仅仅威胁就足够了. 在真实的实验中, 惩罚的确可以提升合作, 但有时候会非常昂贵, 以致于至少在最初的几回合公共利益博弈中, 这类群体的平均收益要低于合作度较低但没有惩罚的群体的平均收益. 而且, 在许多社会中, 不合群的惩罚 (空想社会改良家的惩罚) 是频繁的, 这反而阻碍了惩罚对合作的促进.

1.16 自愿参与

即使声誉可以保持一定数量的惩罚者, 这依然不能解释为什么惩罚行为会出现. 在一个充满背叛者的世界里, 惩罚者需要不停地进行惩罚, 其收益会相当低而使这样的行为不可能流行.

如果参与者不是必须参加公共利益博弈, 而是可以选择退出, 那么情况将有所不同. 这个假设非常自然, 很多时候个体可以决定是否参加集体活动. 在城镇中, 你可以不使用公共交通, 走路也不错. 在一个狩猎采集的部落, 你也可以不参加大型的围猎, 如果你怀疑其他参与者都在怠工. 收集蘑菇或水果也许是一个不错的选择, 你不需要他人协助就可以独立完成.

这样, 即除了参加合作行动的选择之外, 还有一个结果不依赖于他人行为的选择, 我们将发现这将给公共利益博弈带来新的转机. 一个猛犸象进入了你生活的山谷, 你是否会加入捕猎? 加入捕猎可以认为是对合作前景持乐观态度. 我们假设每个参与者都合作, 那么参加公共利益博弈的收益要高于不参加博弈的收益; 但如果其他参与者不合作, 那么参与这类博弈就是浪费时间, 应该选择退出. 此时, 采集蘑菇更合适.

现在考虑这类没有惩罚的 “自愿公共利益博弈”. 此时, 三种策略 (合作、背叛、退出) 以循环方式相互取代, 就如石头 — 剪刀 — 布策略. 如果群体中大部分人在集体行动中选择合作, 那么从利己的角度, 应该选择背叛. 如果大多数人都背叛, 那应该选择退出, 如果多数人不参与博弈, 则应该选择合作. 最后这点看起来有点令人吃惊. 如果少数人愿意参与, 那么多数团队规模都较小. 此时, 虽然在每个团队

内部, 合作者的收益要低于背叛者, 但整体上, 合作者的平均收益要高于背叛者的平均收益.

石头—剪刀—布循环, 从贡献到背叛到退出再到贡献, 不足以建立合作. 从长期来看, 博弈参加者的平均收益并不比博弈退出者的高. 在第 6 章将看到, 如果可以对剥削者进行惩罚, 那么在大多数情况下合作将达成. 这是一个统计结果. 在随机波动条件下, 惩罚者会被二阶剥削者所取代, 后者又将被背叛者所取代, 但随着合作进入低谷, 惩罚者将重新出现. 从长期来看, 他们将主导整个群体.

结果看似矛盾. 在可以自由退出时, 基于胁迫的合作可以在公共利益博弈中出现并流行, 但如果强制参加, 则胁迫会失效而背叛者将获胜.

许多经济学实验证实了这一理论结果的正确性. 在囚徒困境和公共利益博弈中, 如果可以自由选择是否参与博弈, 则合作更容易达成. 有一个特别令人信服的实验, 在一个规模庞大的群体中, 每个回合博弈后参与者有机会选择是否在后续博弈中引入惩罚, 而不是选择是否退出公共利益博弈. 在初始回合, 大多数参与者选择没有惩罚的博弈. 这是可以理解的, 因为没有人希望受到惩罚, 而且很多人也不喜欢惩罚别人. 但五六个回合之后, 几乎所有的参与者都转变了立场, 自愿选择有惩罚的博弈, 进行勤勉的合作. 他们有效地 "用脚投票"(voted with their feet) 来选择惩罚的威胁, 因为他们认为这会使合作的可能性更大. 这个实验看起来就像一个道德剧, 揭示了社会契约的哲学起源.

1.17 艺术的抽象

经济学实验的现实意义如何? 通常最引人注目之处就在于明显的人为痕迹, 而远离日常生活的经验.

这本身并不是一个缺点. 许多经典物理或生理实验同样远离日常生活. 他们的目标是探索自然, 而不是去模仿它. 例如, 可以认为最后通牒博弈的主要优点在于其构建了参与者们从未遇到的情况. 我们总是处于讨价还价, 程式化议价和市场竞争的状态. 最后通牒博弈中的 "要么接受要么放弃" 的无趣选择对大多数参与者而言是完全陌生的. 在没有任何准备的情况下, 实验者迫使我们自发地决定, 而不是依赖习惯的力量.

大多数经济实验以匿名方式进行, 排除了所有可能的影响, 譬如亲缘性, 声誉, 未来的再次交互, 以及广告. 匿名性不是人类进化历程中经常遇到的前提. 人类的进化过程通常发生在小的部落村庄, 人们彼此相互了解. 在进化历史上, 我们不会在不自然的规则下将少量资金转给陌生人. 假设非现实的最后通牒博弈或信任博弈塑造了进化史是没有意义的. 但人类行为是基于进化的特征, 通过在经济学实验中改变一些规则, 我们希望能够揭示这些特征.

例如, 在进行囚徒困境博弈之前, 如果参与者可以进行简短的交流, 那么他们会有更高的合作度. 此外, 通过短暂的谈话, 他们可以非常准确地预测对手是否选择合作. 甚至在不知道具体的博弈实验情况下, 他们也能很快找到相关线索来评估出对手的合作意向. 实验组织者希望通过改变参与者的交流的方式, 譬如面对面, 通过监视器, 通过电话, 或仅仅是一个短暂的视觉接触, 来了解人类是如何评估陌生人的.

正如我们所看到的另一个例子, 如果得知自己的选择将会在未来的合作伙伴中公开, 那么参与者更倾向于拒绝不公平的报价. 然而, 即使确保没有人知道他们的选择, 参与者还是会拒绝不公平的报价. 尽管参与者确信没有人在观察, 并且了解双盲实验的细节, 他们的潜意识仍然可能存在一些疑虑, 而忽略这一事实是幼稚的. 我们还远没有完全理解何时以及为何潜意识会影响决策. 参与者可能会对一些不真实的暗示给予强烈的回应. 由色情杂志引起的性唤起是一个经常被提及的例子.

当然, 实验博弈理论学家知道这一点. 他们不试图重现现实生活中的交互活动以及众多的心理和文化影响, 而旨在分析决策情形的细节. 大多数经济互动在无数附加条件下进行, 并受限于个人经历, 道德传统以及文化规范. 实验必须在这些因素的基础之上进行抽象.

1.18 人类行为的动机

具有自我局限的类似特征, 本书大部分数学模型仅考虑利己而忽略了其他所有心理学因素. 这本身不是严格的限制: 大多数心理学家认为, 这是一个很好的一次近似. (引用乔纳森 · 海德特 (Jonathan Haidt): “如果要你在两秒或者更少的时间内解释人类行为, 你可能想说 ‘利己’.”) 一些有趣且可信的理论假定个人效用包含了其他参与者的效用. 譬如, 对公正的追求, 但是这一重大的心理学问题, 在此不作讨论. 看起来我们的偏好是通过进化产生的, 有一条引导人类由自私的基因通往人性关爱的道路, 但这样的论题超出了本书的范围. 本书仅仅是通过数学手段探索: 利己是如何克服社会困境的.

我们不认可这样的观点, 即我们的社会活动是由每个人内心所固有的 “经济人” 法则来支配的, 人们会通过冷酷的理性计算来选择策略, 使得自身利益达到最大化. 依据约翰 · 梅纳德 · 凯恩斯 (John Maynard Keynes) 的观点, 经济决策往往是由 “动物精神和自发乐观” 来支配的, 并受到 “神经紧张和狂热情绪, 乃至消化功能以及天气状况的影响”.

情感和本能作为启发式的系统通过计算引导我们, 这类计算对我们的理性思维能力造成了严重的负荷. 类似地, 网球运动员试图去估计的轨迹球的速度和精度, 这种能力是任何机器人所无法比拟的. 运动员的计算是在潜意识下完成的, 任何温

布尔顿冠军是否可以通过物理课程的学习而变得更加出色, 这是令人怀疑的. 同样地, 我们不需要用纸和笔来计算我们在实际生活中的得失. 博弈论就像物理学课程, 它有助于我们理解问题, 但却没有给出取得成功的办法.

下面我们将只关注把利己因素作为行为动机的出发点, 而有意忽略社会结构的影响. 与忽略利他主义相比, 忽略社会结构可能是对现实生活一个更为严重的失真. 在篇幅较短的最后一章 (第 7 章) 中简要讨论了一些前述章节所忽略的因素: 家庭关系、邻里效应和群体利益.

因此, 本书的主要部分考虑了在融合良好且规模较大的群体中, 利己个体进行简单的合作博弈. 这确实是一个虚拟场景, 但是我们的世界似乎正朝着它进化发展. 它是未来的样子吗? 它肯定不是我们进化历史的一部分. 我们并未对大城市的生活事先有所准备, 但我们有一种不可思议的与陌生人融合的天赋, 并且能享受 "人头攒动吵吵闹闹的世界", 就像埃德加 · 爱伦 · 坡 (Edgar Allan Poe) 笔下的短篇小说《人群中的人》(*The Man of the Crowd*) 里面的无名英雄一样.

1.19 食物与道德

大多数经济实验通过金钱来完成, 一些批评人士说, 这是仅有的真实性所在. 但事实上, 大量的日常经济合作完全不涉及金钱. 我们可以伸出援手, 或者提供信息, 或者共享一餐: 在每种情况下的心理感受是不同的. 在实验中使用金钱, 是一种简化方法, 它削弱了策略问题复杂化的框架效应.

然而, 很明显, 有时这种结果的标准化可能会带来严重的误导. 例如, 当你考虑 1.3 节中的捐赠博弈时, 有一个场景你可能感到不自在. 如果其他参与者信任你, 你还会背叛吗? 大多数人会在这一点上很犹豫. 通常让别人失望是件令人不愉快的事情. 与这种不愉快的体验相比, 因背叛而额外多得的收益完全不值得. 事实上, 许多实际的实验表明大多数参与者都愿意合作. 相当多的人本性是善良的. 但是这个善良本性来自哪里呢?

在最后通牒博弈中也存在类似问题. 大多数建议者提出分配近一半的金额给对方, 并声称这样才公平. 可以想象, 他们有点自欺欺人, 他们其实只是担心如果提供更低的报价可能会被拒绝. 但为什么接受者会拒绝小的报价呢? 大多数接受者声称他们被明显的不公平报价所激怒. 再一次地, 他们自己混淆了动机, 其实只是想避免导致懦弱的声誉, 这将在长期过程中使他们的收益受损. 然而对于最后通牒博弈的一个变体, 即独裁者博弈, 则对利己的责难就不再成立. 在这种博弈中, 提议者提出分配方案, 接受者不能拒绝, "独裁者" 可以随心所欲.

在独裁者博弈中, 提议者的报价通常低于最后通牒博弈. 而大量的提议者, 给予了对方正的报价. 在这种情况下, 看起来还是很难放弃善良的禀性. 提议者只

是觉得这样做可以表现得慷慨点, 使自己愉快. 如果在变体的独裁者博弈中, 接受者虽然不能拒绝, 但可以写一张小纸条让提议者[1]知道他们的想法, 那么报价会提升至与最后通牒博弈几乎相同的水平. 显然, 人们不愿意招致别人的愤怒, 即使这种愤怒是完全无效的. 此时, 谴责——这类纯符号性的惩罚是几乎没有成本的.

我们只是害怕被诅咒吗? 普遍认为, 人们保持合作和遵守道德是源于对超自然神灵惩罚的恐惧. 由迷信引起的不适应普遍存在, 也许是因为其强烈地促成因循守旧和顺从, 而这些特征通常具有生存价值.

我们享受性爱和食物, 是因为其促进了我们的生存和繁殖. 同样, 我们的生存和繁衍也有赖于我们都是成功的合作者. 这就是为什么我们会选择善良, 为什么我们会觉得报复令人舒畅. 道德情操, 感受善良所产生的内在温暖, 报复非正义所得到的舒畅快意, 过失所引起的内疚羞愧, 这些都深深地根植于我们的天性之中. 文化不同则道德规范不同, 但青少年接受和学习主流规范的能力是他们的天性, 就像接受和学习社区的语言一样.

并不是所有的道德都会促进利他主义和合作. 个人整洁和卫生有类似的道德地位, 但却没有经济学的背景. 很大一部分道德规范最终都归结为促进不断地给予和获得, 这些给予和获得是人类行为的一个重要组成部分. 德国剧作家布莱希特 (Bertold Brecht) 在他的歌剧《三便士》(*Three Penny*) 中写道: "食物第一, 然后道德". 其实恰恰相反. 没有道德, 我们无法生存.

但幸运的是, 我们不需要以训诫说教结束. 此处考虑的简单模型也包含了一些极具价值的教训: 例如, 基于不悦的报复本能, 可以起到阻止背叛者的经济性作用; 或者, 那种一旦可以逃脱惩罚就想剥削别人的自私冲动, 有助于在群体中维持一定数量的报复者, 从而提升共同利益, 又或者在团队合作的前景暗淡时, 选择退出, 实际上有助于强化团队合作. 我们不断发现正是个人所呈现出的不够完美的特征维持了社会凝聚力. 因此, 即便你对利己的本能总是不够满意, 但一想到这有利于公共利益, 大可自我宽慰.

顺便说一句, 人类和太过人性的弱点及错误维持合作是公认的. 这被称为曼德维尔 (Mandeville) 悖论.《蜜蜂寓言》(*Fable of the Bees*) 的作者将下句口号作为副标题: "个人之恶, 公共之益". 利己可促成公共利益. "无形的手" 展示了令人惊讶的技巧.

1.20 参考文献

Hamilton (1996) 和 Trivers (2002) 介绍了合作的进化生物学的基本内容, 也

[1] 译注: 原文此处为 "响应者".

可参见 Trivers (1985), Frank (1998) 以及 Nowak (2006a). 普及工作见 Dawkins (1989), Sigmund (1995) 以及 Ridley (1997). 关于最少技巧的社会困境的博弈理论描述可见 Colman (1995), Binmore (1994), Sugden (1986), Ostrom (1990) 以及 Skyrms (2004). 社会困境的调查方面的优秀工作可参见 Dawes (1980), Cross 和 Guyer (1980), Heckathorn (1996), Kollock (1998), Levin (1999). Hardin (1968) 以及 Olson (1965) 研究了公共利益悲剧以及围绕集体行动的困境问题. 关于囚徒困境博弈的流行解释由 Poundstone (1992) 提出. 囚徒困境博弈的概念由 Luce 和 Raiffa (1957) 首次提出, 也可参见 Schelling (1978). Rapoport 和 Chammah (1965) 为专门论述博弈论的著作. Rapoport 参加计算机排名赛 (Axelrod 1984) 并提交了以牙还牙策略. 间接回报最早由 Alexander (1987) 和 Ellison (1994) 提出. Nowak 和 Sigmund (1998a, 1998b) 建立了相关模型, Wedekind 和 Milinski (2000) 以及 Wedekind 和 Braithwaite (2002) 作了早期的实验. 雪堆博弈由 Sugden (1986) 提出, 也可参见 Doebeli, Hauert 和 Killingback (2004). 信任博弈最早由 Berg, Dickhaut 和 McCabe (1995) 提出. 最后通牒博弈则由 Güth, Schmittberger 和 Schwarze (1982) 提出. Henrich 等 (2006) 利用最后通牒博弈研究了小型团体. Kreps 和 Wilson (1982) 以及 Kurzban, DeScioli 和 O'Brien (2007) 研究了声誉在经济学中的作用. Haley 和 Fessler (2005), Bateson, Nettle 和 Roberts (2006) 以及 Burnham 和 Hare (2007) 研究了观察者的作用. Yamagishi (1986) 以及 Fehr 和 G?chter (2000, 2002) 研究了惩罚在公共利益博弈中的作用, 这类工作也可见 Boyd 和 Richerson (1992), O'Gorman, Wilson 和 Miller (2005), Gardner 和 West (2004) 以及 Sigmund (2007). Herrmann, Thöni 和 Gächter (2008) 研究了反社会的惩罚中比较困难的问题. Gürerk, Irlenbusch 和 Rockenbach (2006) 通过实验发现只有在经历几个回合的博弈之后, 公共利益博弈的参与者才更倾向于惩罚背叛者. 关于正向及负向激励对于促进合作所起的作用, 许多学者作了对比研究, 参见 Baumeister 等 (2001), Dickinson (2001), Andreoni, Harbaugh 和 Vesterlund (2003), Walker 和 Halloran (2004) 以及 Sefton, Shupp 和 Walker (2007). Orbell 和 Dawes (1993), Hauert 等 (2002a, 2002b) 以及 Fowler (2005a) 研究了自愿参与的作用. Kagel 和 Roth (1995) 以及 Camerer (2003) 这两部专著全面介绍了经济学实验以及行为博弈, 也可参见 Camerer 和 Fehr (2006). Carpenter, Harrison 和 List (2005) 介绍了规则更为自然的实验. 在理论层面上, Yamagishi, Jin 和 Kiyonari (1999), Bendor 和 Swistak (2001), Price, Cosmides 和 Tooby (2002), Ostrom 和 Walker (2003), Cose (2004), Kurzban 和 Houser (2005), Hauser (2006) 以及 Haidt (2007) 从心理学和社会学的角度研究了伦理准则和道德规范. Brown (1991) 讨论了人类的共性. Richerson 和 Boyd (2005) 在文化进化和生物进化方面作了统一论述. Dunbar (1996) 以及 Sommerfeld 等 (2007) 揭示了闲聊的作用. 关于感性承诺在经济学中的作用, Frank (1988) 是非常重要的文献. 关于经济

博弈相关的情绪, Rilling 等 (2002), Sanfey 等 (2003), de Quervain 等 (2004) 研究了其在神经系统方面的基础. Charness 和 Haruvy (2002) 以及 Sachs 等 (2004) 指出了回报的重要性. Bowles 和 Gintis (2002) 以及 Gintis 等 (2003, 2005) 提出了具有重要意义的方法——“强回报”: 其中核心观点可见 Burnham 和 Johnson (2005). Fehr 和 Schmidt (1999) 以及 Bolton 和 Ockenfels (2000) 阐述了如何通过修改效用解释实验结果, 以达到对公平和正义的追求. 大量文献研究了非人灵长类动物的社会交互, 可参见 de Waal (1996), Brosnan 和 de Waal (2003), Stevens, Cushman 和 Hauser (2005), Silk (2006) 以及 Warneken 和 Tomasello (2006). Clutton-Brock 和 Parker (1995), Kiers 等 (2003) 以及 Wenseleers 和 Ratnieks (2006) 讨论了动物群体内各种形式的惩罚, 关于其他抑制竞争的方法, 可见 Frank (2003).

1.21 博弈小结: 二人博弈的简明词典

许多二人博弈实验都是围绕社会困境展开的. 通常参与者匿名, 且事先被赠与了一笔钱 (例如, 出场费). 在熟悉规则且被指定为提议者或响应者 (捐赠者或接收者) 之后, 他们被要求做出选择.

捐赠 (donation) 博弈: 在某种意义上, 它是一种原子型的社会交互. 捐赠者决定是否付出一元, 而让接收者获得三元的收益.

囚徒困境博弈: 各类变种的合作博弈的母版. 一个特别易懂的规则是: 两位参与者相互进行捐赠博弈. 当参与者同时做出选择时, 其与公共利益博弈非常类似. 如果都选择合作 (即捐赠), 那么双方都会获得收益两元. 因为捐赠意味着需支付一元, 所以无论对手如何选择, 最佳的做法是拒绝捐赠 (即背叛). 如果两位参与者都背叛, 则他们无收益.

最后通牒博弈: 实验组织者提供一笔钱, 提议者提出分一定金额 (报价) 给响应者. 如果响应者 (知道这笔钱的总额) 接受, 则据此分割, 博弈结束. 否则, 如果响应者拒绝, 实验组织者收回那笔钱, 两位参与者一无所获.

独裁者博弈: 除了响应者不能拒绝报价这点, 与最后通牒博弈完全相同.

信任博弈: 第一阶段, 如捐赠博弈那样, 提议者 (投资者) 给响应者 (受托人) 一笔钱. 第二阶段, 响应者决定返还提议者一部分钱. 这与串行化的囚徒困境博弈 (即首先一位参与者扮演捐赠者的角色, 然后另一位再扮演) 类似.

重复囚徒困境博弈: 两位参与者进行多个回合的囚徒困境博弈. 通常, 他们事先不知道博弈什么时候结束, 以避开 “最后回合效应” (因对手不能在单回合囚徒困境博弈中进行报复而诱导的背叛).

间接回报: 在规模较大的群体中, 随机抽取两位参与者进行捐赠博弈或者单回合囚徒困境博弈. 重复这一过程. 参与者知道只相遇一次, 对手没有机会报复.

雪堆博弈: 两位参与者事先向实验组织者支付一笔费用, 以获得一笔钱 (高于支付的费用). 参与者必须决定是否支付费用, 注意如果都愿意支付, 则每位参与者只需各承担一半.

第 2 章　博弈动力学与社会性学习

2.1　博　　弈

我们很难决定哪种策略是最好的. 因为事件进程充满了不确定性 (就像一个投资者想要最优的投资策略一样), 或者确定事件进程需要大量复杂计算 (正如一个旅行商要找到一条通过 87 个城市的最短路径). 如果两个以上的决策者对最优策略的观点不一致, 则会产生特别复杂的情况. 这属于博弈论的研究领域.

举个例子, 考虑参与者 I 和 II 进行如下儿童游戏. 两人同时伸出一根或两根手指. 如果双方手指数之和是奇数, 则参与者 I 获胜, 如果是偶数, 则参与者 II 获胜. 两人不能同时获胜. 其中一人总是会倾向切换到另一个选择. 这种情形普遍存在于社会交往中.

让我们将这类问题规范化. 假设参与者 I 有 n 种**策略**(strategy), 我们用 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 表示, 参与者 II 有 m 种策略, 用 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_m$ 表示. 如果 I 选择 $\boldsymbol{e}_i$ 而 II 选择 $\boldsymbol{f}_j$, 那么 I 的**收益**(payoff) 为 a_{ij}, II 的收益为 b_{ij}. 这个博弈可以用两个 $n \times m$ 的收益矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 表示, 或者用一个 i 行 j 列元素为一对收益值 (a_{ij}, b_{ij}) 的矩阵表示. 衡量收益的效用尺度是由参与者的偏好决定的. 在一个生物博弈中, 衡量繁殖成功率的达尔文适合度可以看作一种效用尺度. 为了简单起见, 我们选择货币来衡量收益.

在上面的例子中, 假设两个参与者的赌注是 1 元. 每个参与者有两种策略, 伸出**偶数** (even) 根或**奇数** (odd) 根手指, 对于 I 我们分别用 $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ 来表示, 对于 II 我们分别用 $\boldsymbol{f}_1$, $\boldsymbol{f}_2$ 来表示. 那么它的收益矩阵是

$$\begin{pmatrix} (-1,1) & (1,-1) \\ (1,-1) & (-1,1) \end{pmatrix}. \tag{2.1}$$

如果结果是 $(-1,1)$, 参与者 I(策略对应矩阵行的参与者) 选另一种策略会更好; 如果结果是 $(1,-1)$, 参与者 II (策略对应矩阵列的参与者) 选另一种策略会更好. 如果参与者能看透彼此, 他们会陷入一个恶性循环.

2.2　混 合 策 略

对两个参与者来说, 不被对手预测到自己的决定是最重要的. 能达到这种效果的一个好方法就是随机, 即凭运气去决定. 假设参与者 I 选择策略 $\boldsymbol{e}_i$ 的概率是 x_i,

则混合策略 (mixed strategy) 可以用一个随机向量 $\boldsymbol{x}=(x_1,\cdots,x_n)$ 来表示 (其中 $x_i\geqslant 0, x_1+\cdots+x_n=1$). 我们记所有这些混合策略的集合为 S_n: 这是 $\mathbf{R}^n$ 上的一个单形, 它是由标准基当中的单位向量 $\boldsymbol{e}_i$ 张成的, 单位向量 $\boldsymbol{e}_i$ 称为纯策略 (pure strategy), 对应于原有策略集合中的单个策略, 如图 2.1 所示, ($\boldsymbol{e}_i$ 除了第 i 个分量为 1, 其余分量都为 0.)

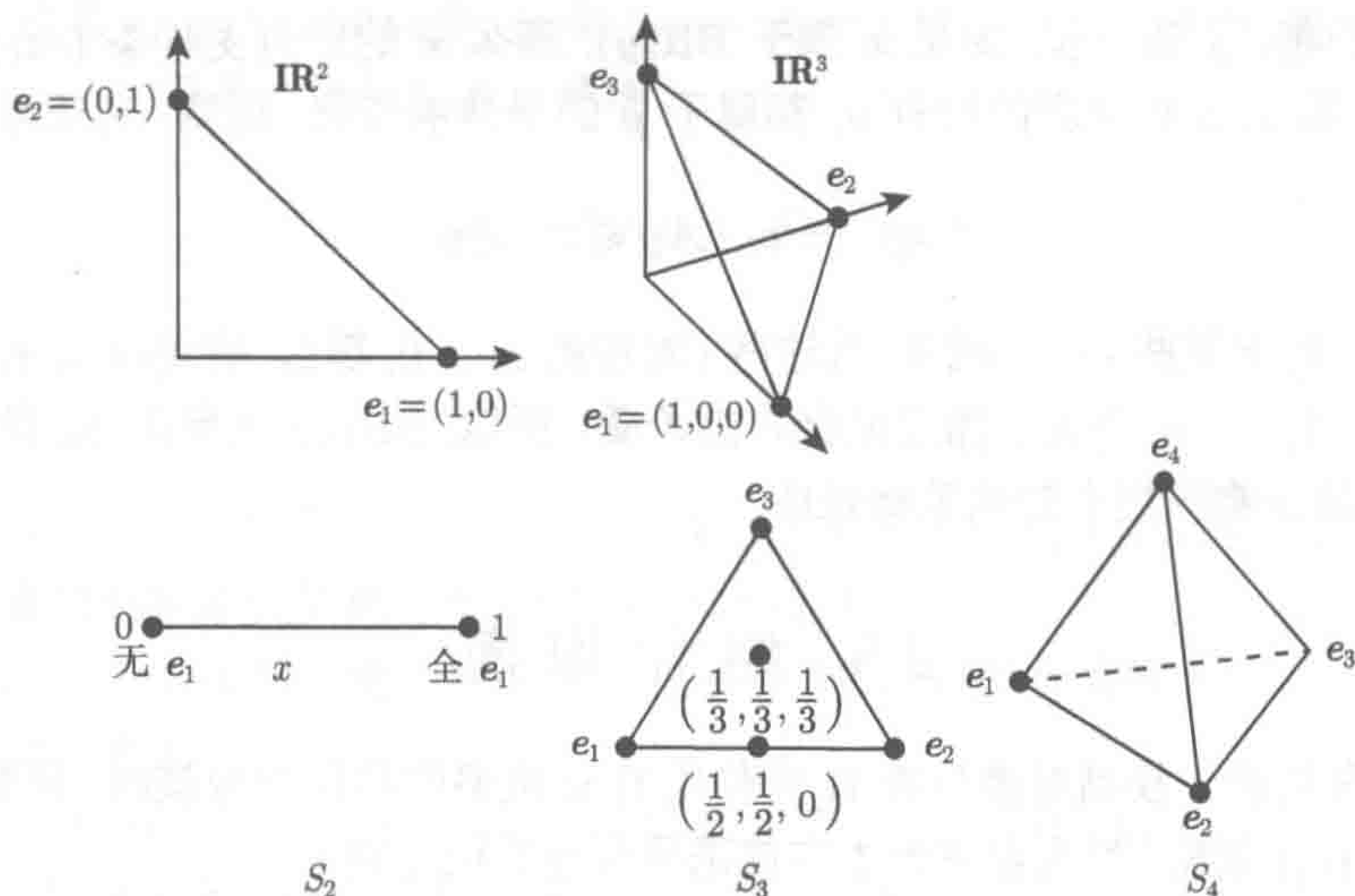


图 2.1 单形 S_2, S_3 和 S_4. 在第一行, S_2 和 S_3 分别嵌入到 $\mathbf{R^2}$ 和 $\mathbf{R^3}$ 中

类似地, 参与者 II 的混合策略可以用单位单形 S_m 的一个元素 $\boldsymbol{y}$ 表示. 如果参与者 I 使用纯策略 $\boldsymbol{e}_i$ 且参与者 II 使用策略 $\boldsymbol{y}$, 那么参与者 I 的收益 (或者更精确地讲, 它的期望值) 为

$$(\boldsymbol{A}\boldsymbol{y})_i=\sum_{j=1}^{m}a_{ij}y_j. \tag{2.2}$$

如果参与者 I 用混合策略 $\boldsymbol{x}$, 参与者 II 用混合策略 $\boldsymbol{y}$, 那么参与者 I 的收益为

$$\boldsymbol{x}\cdot\boldsymbol{A}\boldsymbol{y}=\sum_i x_i(\boldsymbol{A}\boldsymbol{y})_i=\sum_{i,j}a_{ij}x_iy_j, \tag{2.3}$$

类似地, 参与者 II 的收益为

$$\boldsymbol{x}\cdot\boldsymbol{B}\boldsymbol{y}=\sum_{i,j}b_{ij}x_iy_j. \tag{2.4}$$

(左边的点表示点积或欧几里得积)

如果参与者 I 碰巧知道了对方会使用策略 $\boldsymbol{y}$, 那么参与者 I 应该会使用对 $\boldsymbol{y}$ 来说的最佳响应策略 (best reply strategy). 最佳响应策略组成的集合可以表示为

$$\mathrm{BR}(\boldsymbol{y})=\arg\max_{\boldsymbol{x}}\boldsymbol{x}\cdot\boldsymbol{A}\boldsymbol{y}, \tag{2.5}$$

即, 所有的 $\boldsymbol{x} \in S_n$, 使得对任意的 $\boldsymbol{z} \in S_n$,

$$\boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{y} \leqslant \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{y} \tag{2.6}$$

都成立. 参与者 I 没有动机偏离 $\boldsymbol{x}$ 这个策略.

因为函数 $\boldsymbol{z} \mapsto \boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{y}$ 是连续的且 S_n 是紧的, 所以最佳响应集合总是非空的. 它是一个凸集. 更进一步, 如果 $\boldsymbol{x}$ 属于 $\mathrm{BR}(\boldsymbol{y})$, 那么 $\boldsymbol{x}$ 的所有**支撑集** (support) 中的纯策略, 即 $x_i > 0$ 对应的所有 $\boldsymbol{e}_i$ 都属于 $\boldsymbol{y}$ 的最佳响应集. 确实, 对任意的 i 有

$$(\boldsymbol{A}\boldsymbol{y})_i = \boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{y} \leqslant \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{y}. \tag{2.7}$$

如果不等号对于某些 i 严格成立, 且这些 i 对应的 $x_i > 0$, 那么 $x_i(\boldsymbol{A}\boldsymbol{y})_i < x_i(\boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{y})$; 将所有 $i = 1, \cdots, n$ 个式子加起来会导出矛盾. 所以, $\mathrm{BR}(\boldsymbol{y})$ 是单形 S_n 的一个面. 它由 $\boldsymbol{y}$ 的最佳响应集中的纯策略张成.

2.3 纳什均衡

如果参与者 I 找到对参与者 II 所使用的 $\boldsymbol{y}$ 策略的最佳响应策略, 只要参与者 II 坚持使用 $\boldsymbol{y}$ 策略, 那么参与者 I 应该选择最佳响应策略.

但是参与者 II 会坚持选择 $\boldsymbol{y}$ 策略吗? 只有在参与者 II 没有动机选择其他策略的情况下才会发生, 即, II 也恰好采用了最佳响应策略. 如果 $\boldsymbol{x}, \boldsymbol{y}$ 互为对方的最佳响应策略, 那么这两个策略被称作一个**纳什均衡对** (Nash equilibrium pair), 即, 如果 $\boldsymbol{x} \in \mathrm{BR}(\boldsymbol{y})$ 且 $\boldsymbol{y} \in \mathrm{BR}(\boldsymbol{x})$, 换句话说, 如果

$$\boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{y} \leqslant \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{y} \tag{2.8}$$

对所有的 $\boldsymbol{z} \in S_n$ 成立, 且

$$\boldsymbol{x} \cdot \boldsymbol{B}\boldsymbol{w} \leqslant \boldsymbol{x} \cdot \boldsymbol{B}\boldsymbol{y} \tag{2.9}$$

对所有的 $\boldsymbol{w} \in S_m$ 成立. 纳什均衡对 $(\boldsymbol{x}, \boldsymbol{y})$ 满足最小一致性要求: 参与者没有动机偏离 (只要其他参与者不偏离).

一个基本的结果: 任何博弈 $(\boldsymbol{A}, \boldsymbol{B})$ 都存在一个纳什均衡对, 这将在 2.11 节给出证明. 这个结论对任意数量的参与者, 对任意策略的紧凸集, 对任意连续的收益函数, 甚至对更广泛的情形都成立. 如果不允许使用混合策略的话, 博弈可能不存在纳什均衡: 这从 2.1 节的简单例子可以看出 (赌奇偶). 在那种情况下, 选择奇或偶的概率都是 1/2 的混合策略显然是一个纳什均衡对: 平均每一个参与者都获得零元, 而且没有人有动机偏离这个策略. 另一方面, 如果参与者 I 采用任何其它策略 (x_1, x_2) 应对参与者 II 的策略 $(1/2, 1/2)$, 参与者 I 的收益仍然是零元. 然而, 另

一个参与者就有动机偏离了: 当 $x_1 > x_2$ 时, 参与者 II 的最佳响应策略是 (1,0). 如果参与者 II 使用了这个策略, 那么参与者 I 最好使用策略 (0,1), 这就陷入了一个全面混乱的死循环.

在这个例子中, 当 $\boldsymbol{x} = \boldsymbol{y} = (1/2, 1/2)$ 时, 混合策略 $(\boldsymbol{x}, \boldsymbol{y})$ 是唯一的纳什均衡对. 我们已经知道, 只要 II 选择均衡策略 $\boldsymbol{y}$, 那么 I 就没有理由偏离均衡策略 $\boldsymbol{x}$; 但如果 II 不选策略 $\boldsymbol{y}$, I 就会改变. 如果 $(\boldsymbol{x}, \boldsymbol{y})$ 是一个严格纳什均衡对, 那么情况会有所不同. $(\boldsymbol{x}, \boldsymbol{y})$ 是严格纳什均衡对是指

$$\boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{y} < \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{y} \tag{2.10}$$

对所有的 $\boldsymbol{z} \neq \boldsymbol{x}$ 都成立, 且

$$\boldsymbol{x} \cdot \boldsymbol{B}\boldsymbol{w} < \boldsymbol{x} \cdot \boldsymbol{B}\boldsymbol{y} \tag{2.11}$$

对所有的 $\boldsymbol{w} \neq \boldsymbol{y}$ 都成立. 在这种情况下, 即当所有的最佳响应集都是单点集时, 每一个参与者如果单方面偏离均衡策略都会受到惩罚.

虽然每一个博弈都有纳什均衡对, 但并非都有严格纳什均衡对, 2.1 节的数字游戏就是一个例子.

而且, 即使存在严格纳什均衡对, 结果也有可能让人失望, 就好像 1.3 节囚徒困境的例子. 此博弈有一个严格纳什均衡对: 参与者都背叛, 即 $\boldsymbol{x} = \boldsymbol{y} = (0, 1)$. 任何一方改变策略都将受到惩罚. 然而, 如果双方都偏离均衡, 选择合作的话, 他们都会更好.

值得注意的是: 对很多博弈来说, 可能存在不止一个纳什均衡对. 参与者会选择哪一个呢? 当然, 他们可以坐下来谈一谈, 但是我们不考虑这种情况. 大部分情况参与者不能交流, 有时交流是被明令禁止的, 有时交流只是在浪费口舌. 考虑 1.4 节的雪堆问题, 很容易看出来, 它有两个纳什均衡对 $(\boldsymbol{e}_1, \boldsymbol{f}_2)$ 和 $(\boldsymbol{e}_2, \boldsymbol{f}_1)$. 在旁观者眼里这两种状态是没有差别的, 但是参与者的感受不同. 策略对 $(\boldsymbol{e}_1, \boldsymbol{f}_2)$ 意味着 I 铲雪而 II 躺下休息. I 不愿意这样, 但却不能不去铲雪, 因为不去铲雪意味着要在车里过一夜. 当然 I 期望第二种纳什均衡对, 为此 I 选择了 $\boldsymbol{e}_2$, 而对 II 来说第一种纳什均衡对更好, 如果他顽固地选 $\boldsymbol{f}_2$, 那么最终会导致出现了策略对 $(\boldsymbol{e}_2, \boldsymbol{f}_2)$ (在车里度过不舒服而又寒冷的一晚), 这不是一个纳什均衡对.

2.4 群体博弈

到目前为止, 考虑的博弈都是在两个具体的参与者之间进行的, 且他们尝试猜测对方的策略, 并找到一个最佳响应策略. 让我们转换一下视角, 考虑一个参与者**群体** (population), 每一个参与者都有一个固定策略. 不时地, 随机让两个参与者相

遇并用其固定策略进行博弈. 将这些策略作为一种行为模式来考虑. 这种模式可能被学习、遗传或以任何其它方式打上烙印. 在生态环境里, 这些策略就好像是不同类型的个体 (或者说行为表型).

为了分析这种结构, 我们简单的假设群体中的个体只是用相互作用的方式来予以区分, 即参与者只有类型或者说使用策略不同. 这个很好地适用于囚徒困境这一类的博弈, 此时, 参与者信息是对称的; 但是在社会交往的许多例子中, 存在固有的不对称 —— 例如, 在买家和卖家之间.

为了简化起见, 我们先考虑**对称博弈** (symmetric game). 在两个参与者的博弈中, 这意味着如果参与者 I 和参与者 II 互换, 博弈仍然不会改变. 特别地, 当两个参与者有相同的策略集合的时候. 因此, 我们假设 $n = m$, 且对所有的 j, $\boldsymbol{f}_j = \boldsymbol{e}_j$; 如果参与者选择 $\boldsymbol{e}_i$ 策略应对使用 $\boldsymbol{e}_j$ (以前是策略 $\boldsymbol{f}_j$) 的对手, 那么他被标记成 I 或者 II 都会有相同的收益. 因此 $a_{ij} = b_{ji}$, 换句话说 $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ (矩阵 $\boldsymbol{A}$ 的转置). 所以对称博弈可以用矩阵对 $(\boldsymbol{A}, \boldsymbol{A}^{\mathrm{T}})$ 表示, 因而可以简单地由收益方阵 $\boldsymbol{A}$ 来定义.

对称博弈可以有非对称的纳什均衡对, 正如我们已经看到的雪堆博弈. 如果不能区分参与者 I 和 II, 那么这些显然无关紧要. 我们感兴趣的只是对称纳什均衡对, 即当 $\boldsymbol{x} = \boldsymbol{y}$ 时的策略对 $(\boldsymbol{x}, \boldsymbol{y})$. 因此, **对称纳什均衡** (symmetric Nash equilibrium) 可以只用一个策略 $\boldsymbol{x}$ 表示, 且它的最佳响应策略就是它自己 (即 $\boldsymbol{x} \in \boldsymbol{BR}(\boldsymbol{x})$). 换句话说, 它有如下性质:

$$\boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x} \tag{2.12}$$

对任意的 $\boldsymbol{z} \in S_n$ 成立. 一个对称的严格纳什均衡

$$\boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{x} < \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x} \tag{2.13}$$

对任意的 $\boldsymbol{z} \neq \boldsymbol{x}$ 成立.

我们将证明任何对称博弈都有对称纳什均衡.

2.5　博弈对称化

有一个明显的方法可以把非对称博弈 $(\boldsymbol{A}, \boldsymbol{B})$ 变成对称博弈: 用掷硬币的方法决定两个参与者谁被标记成 I. 因此**对称化** (symmetrized) 博弈的策略必须指定角色 I 应该做什么, 角色 II 应该做什么, 即策略是由策略对 $(\boldsymbol{e}_i, \boldsymbol{f}_j)$ 决定的. 混合策略由元素 $\boldsymbol{z} = (z_{ij}) \in S_{nm}$ 给定, 其中 z_{ij} 表示当他是角色 I 时选择 $\boldsymbol{e}_i$, 而当他是角色 II 时选择 $\boldsymbol{f}_j$ 的概率. $\boldsymbol{z}$ 的概率分布符合它的**边际** (marginal): $x_i = \sum_j z_{ij}$ 和 $y_j = \sum_i z_{ij}$. 向量 $\boldsymbol{x} = (x_i)$ 和 $\boldsymbol{y} = (y_j)$ 分别属于 S_n 和 S_m. 容易得到, 任意给定 $\boldsymbol{x} \in S_n$ 和 $\boldsymbol{y} \in S_m$, 存在 $\boldsymbol{z} \in S_{nm}$ 以 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 作为边际, 例如 $z_{ij} = x_i y_j$ (通常, 还有很多边际相同的概率分布).

一方使用策略 $(\boldsymbol{e}_i, \boldsymbol{f}_j)$ 对另一方使用策略 $(\boldsymbol{e}_k, \boldsymbol{f}_l)$ 的收益 $(i, k \in \{1, \cdots, n\}$, $j, l \in \{1, \cdots, m\})$, 依赖于掷硬币的结果, 为

$$c_{ij,kl} = \frac{1}{2}a_{il} + \frac{1}{2}b_{kj}. \tag{2.14}$$

假设对称博弈总有对称的纳什均衡, 则容易推出每个博弈 $(\boldsymbol{A}, \boldsymbol{B})$ 都有纳什均衡对.

如果 $\bar{\boldsymbol{z}} \in S_{nm}$ 是对称博弈 $(\boldsymbol{C}, \boldsymbol{C}^{\mathrm{T}})$ 的一个对称纳什均衡, 则

$$\boldsymbol{z} \cdot \boldsymbol{C}\bar{\boldsymbol{z}} \leqslant \bar{\boldsymbol{z}} \cdot \boldsymbol{C}\bar{\boldsymbol{z}} \tag{2.15}$$

对任意 $\boldsymbol{z} \in S_{nm}$ 成立. 设 $\boldsymbol{z}$ 的边际为 $\boldsymbol{x}$ 和 $\boldsymbol{y}$, $\bar{\boldsymbol{z}}$ 的边际为 $\bar{\boldsymbol{x}}$ 和 $\bar{\boldsymbol{y}}$. 那么

$$\boldsymbol{z} \cdot \boldsymbol{C}\bar{\boldsymbol{z}} = \sum_{ijkl} z_{ij} c_{ij,kl} \bar{z}_{kl} \tag{2.16}$$

$$= \frac{1}{2}\sum_{ijkl} z_{ij} a_{il} \bar{z}_{kl} + \frac{1}{2}\sum_{ijkl} z_{ij} b_{kj} \bar{z}_{kl} \tag{2.17}$$

$$= \frac{1}{2}\sum_{il} x_i a_{il} \bar{y}_l + \frac{1}{2}\sum_{jk} y_j b_{kj} \bar{x}_k = \frac{1}{2}\boldsymbol{x} \cdot \boldsymbol{A}\bar{\boldsymbol{y}} + \frac{1}{2}\boldsymbol{y} \cdot \boldsymbol{B}^{\mathrm{T}}\bar{\boldsymbol{x}}. \tag{2.18}$$

因为 $\bar{\boldsymbol{z}}$ 是一个对称纳什均衡, 所以由不等式 (2.15) 可得

$$\boldsymbol{x} \cdot \boldsymbol{A}\bar{\boldsymbol{y}} + \bar{\boldsymbol{x}} \cdot \boldsymbol{B}\boldsymbol{y} \leqslant \bar{\boldsymbol{x}} \cdot \boldsymbol{A}\bar{\boldsymbol{y}} + \bar{\boldsymbol{x}} \cdot \boldsymbol{B}\bar{\boldsymbol{y}}. \tag{2.19}$$

当 $\boldsymbol{y} = \bar{\boldsymbol{y}}$ 时, 有

$$\boldsymbol{x} \cdot \boldsymbol{A}\bar{\boldsymbol{y}} \leqslant \bar{\boldsymbol{x}} \cdot \boldsymbol{A}\bar{\boldsymbol{y}}, \tag{2.20}$$

当 $\boldsymbol{x} = \bar{\boldsymbol{x}}$ 时, 有

$$\bar{\boldsymbol{x}} \cdot \boldsymbol{B}\boldsymbol{y} \leqslant \bar{\boldsymbol{x}} \cdot \boldsymbol{B}\bar{\boldsymbol{y}}. \tag{2.21}$$

因此, $\bar{\boldsymbol{x}} \in \boldsymbol{B}R(\bar{\boldsymbol{y}})$ 和 $\bar{\boldsymbol{y}} \in \boldsymbol{B}R(\bar{\boldsymbol{x}})$, 即 $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$ 是博弈 $(\boldsymbol{A}, \boldsymbol{B})$ 的纳什均衡对.

2.6 群体动力学与博弈论

现在考虑一个具有收益矩阵 $\boldsymbol{A}$ 的对称博弈, 并且假设在一个大且均匀混合的群体中, 采用策略 $\boldsymbol{e}_i, i = 1, \cdots, n$ 的参与者占参与者总数的比例为 x_i. 群体的状态用向量 $\boldsymbol{x} \in S_n$ 表示, 一个参与者使用策略 $\boldsymbol{e}_i$ 的期望收益为

$$(\boldsymbol{A}\boldsymbol{x})_i = \sum_j a_{ij} x_j. \tag{2.22}$$

实际上, 参与者遇见使用 $\boldsymbol{e}_j$ 策略的对手的概率是 x_j. 群体的平均收益为

$$\boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x} = \sum_i x_i(\boldsymbol{A}\boldsymbol{x})_i. \tag{2.23}$$

应该强调的是这里有一点符号的混淆. 相同的符号 $\boldsymbol{x} \in S_n$, 前面章节表示一个特定参与者的混合策略, 现在表示由不同类型组成的群体的状态, 每种类型采用不同的纯策略. (我们也可以让参与者使用混合策略, 但是现在不需要考虑这种情况.)

如下为基本点: 假设群体是进化的, 这意味着 x_i 将会随时间变化. 令状态 $\boldsymbol{x}(t)$ 依赖于时间, 用 $\dot{x}_i(t)$ 表示 x_i 随时间改变的速率, 即 $\dot{x}_i = \mathrm{d}x_i/\mathrm{d}t$. 按照群体动力学, 我们特别感兴趣的是策略频度的单位增长率 $\dot{x}_i/x_i$.

策略频度是怎么进化的呢? 如何增长和减少的呢? 有很多方法来模拟这个过程. 我们通常假设群体进化状态满足*复制方程* (replicator equation). 如果策略频度的增长率对应策略的收益, 或更准确地, 等于策略收益 $(\boldsymbol{A}\boldsymbol{x})_i$ 和群体平均收益 $\boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x}$ 的差, 就为复制方程. 这样, 我们有

$$\dot{x}_i = x_i[(\boldsymbol{A}\boldsymbol{x})_i - \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x}], \quad i = 1, \cdots, n. \tag{2.24}$$

因此, 策略 $\boldsymbol{e}_i$ 的增长或减少依赖于其比平均更好或更糟.

这样就产生了一个群体状态的确定性模型. 实际上, 任意右端光滑的*常微分方程* (ordinary differential equation) $\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x})$ (比如 (2.24)) 对每一个初值 $\boldsymbol{x}$ 都有唯一解, 即从开区间 I (包含 0) 到 $\mathbf{R}^n$ 的函数 $t \mapsto \boldsymbol{x}(t)$, 使得 $\boldsymbol{x}(0) = \boldsymbol{x}$, 并对任意的 $t \in I$ 有 $\dot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}(t))$ 成立. 本书讨论的所有微分方程, 区间 I 总是取成整个实轴 $\mathbf{R}$.

可以把微分方程右端当作一个向量场 $\boldsymbol{x} \mapsto \boldsymbol{F}(\boldsymbol{x})$. 在 $\boldsymbol{F}$ 的定义域 ($\mathbf{R}^n$ 的一个开子集 B) 中的每一个点 $\boldsymbol{x}$, 都有一个 "风速" $\boldsymbol{F}(\boldsymbol{x}) \in \mathbf{R}^n$. 方程的解描述了在 0 时刻位于 $\boldsymbol{x}$ 点的粒子随风地运动. 在 $\boldsymbol{z}$ 点, 如果 $\boldsymbol{F}(\boldsymbol{z}) = \mathbf{0}$, 则粒子的速度为零. 此时, 对应*驻点* (rest point): 粒子在 $\boldsymbol{z}$ 点是静止的. 注意到在方程右端 $\boldsymbol{F}(\boldsymbol{x})$ 乘以正函数 $M(\boldsymbol{x}) > 0$ 对应*速度的改变* (change in velocity). 粒子将在相同的轨道上以不同的速度移动.

在解释 (2.7 节) 为什么对方程 (2.24) 感兴趣之前, 我们应该注意到 $\sum \dot{x}_i = 0$. 进一步, 我们容易看到对任意的 t, 常函数 $x_i(t) = 0$ 显然满足 (2.24) 的第 i 个方程. 因此, 状态空间, 即单形 S_n 是不变的: 如果 $\boldsymbol{x}(0) \in S_n$, 那么对任意的 $t \in R$ 有 $\boldsymbol{x}(t) \in S_n$. 同样的, 不变性对 S_n 的子单形 ($x_i = 0$, 对某些 i 成立) 也成立. 因此, 对 S_n 的边界 $\mathbf{bd}S_n$ 也成立 (即所有子单形的并集), 同样对单形 S_n 的内部 $\mathrm{int}S_n$ (满足对任意的 i, $x_i > 0$) 也成立.

2.7 模仿动力学

复制方程最初源于生物博弈. 在收益对应于繁殖成功率, 且个体纯育的假设下, 可以立即导出这个方程. 显然, 对于此处考虑的经济学博弈来说, 策略不可能遗传, 但它们可以通过社会学习进行传播. 如果假设个体可以相互模仿, 可以再次得到复制方程.

更准确地讲, 被随机选中的个体从群体中随机找到一个榜样, 然后以某种确定的可能性来模仿这个榜样. 因此, 在间隔长度为 Δt 的时间内, 个体从策略 $\boldsymbol{e}_j$ 转变到 $\boldsymbol{e}_i$ 的概率为 $x_i f_{ij} \Delta t$. 这对应投入产出方程

$$x_i(t+\Delta t) - x_i(t) = \sum_j f_{ij} x_i x_j \Delta t - \sum_j f_{ji} x_i x_j \Delta t, \text{①} \tag{2.25}$$

在 $\Delta t \mapsto 0$ 时有

$$\dot{x}_i = x_i \sum_j (f_{ij} - f_{ji}) x_j. \tag{2.26}$$

一般地, f_{ij} 依赖于状态 $\boldsymbol{x}$. 例如, 我们假设

$$f_{ij} = [(\boldsymbol{Ax})_i - (\boldsymbol{Ax})_j]_+. \tag{2.27}$$

这意味着一个使用 $\boldsymbol{e}_j$ 的参与者通过与使用 $\boldsymbol{e}_i$ 的参与者进行比较, 只有当后者的策略有更好的收益时, 前者才会采取后者的策略: 并且在这种情况下, 收益差值越大, 改变策略的可能性就越大. 此时, 因为 $f_{ij} - f_{ji} = (\boldsymbol{Ax})_i - (\boldsymbol{Ax})_j$, 所以投入产出方程为

$$\dot{x}_i = x_i \sum_j [(\boldsymbol{Ax})_i - (\boldsymbol{Ax})_j] x_j = x_i [(\boldsymbol{Ax})_i - \boldsymbol{x} \cdot \boldsymbol{Ax}], \tag{2.28}$$

这就是复制方程 (2.24). 我们也可以通过类似的方式得到复制方程, 现在用更一般的 “适合度” 来度量策略的成功率, 用以代替收益 $(\boldsymbol{Ax})_i$. 例如 $(1-s)\boldsymbol{B} + s(\boldsymbol{Ax})_i$, $0 < s \leqslant 1$. 这是 “基准适合度” $\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{x}) > 0$ (对所有的类型都适用) 和收益的一个凸组合. s 的大小是博弈中评估策略 “吸引度” 的重要标准.

我们也可以假设

$$f_{ij} = (1-s)\boldsymbol{B} + s(\boldsymbol{Ax})_i, \tag{2.29}$$

这表示比率只依赖于这个榜样的成功率 (而与采用 $\boldsymbol{e}_j$ 策略参与者的收益无关); 或者

$$f_{ij} = (1-s)\boldsymbol{B} - s(\boldsymbol{Ax})_j, \tag{2.30}$$

①译注: 原文此处为 “$\sum f_{ij} x_i x_j \Delta t - \sum f_{ji} x_i x_j \Delta t$”.

这表示参与者越是对自己的收益不满意, 越是倾向于模仿别人. 这个凸组合保证(至少在 s 很小的时候) 比率是正的.

并不是所有的模仿机制都可以导出一个复制方程. 例如, 我们可以假设两个参与者比较他们的收益, 差的总会模仿好的. 因此, 如果 $(\boldsymbol{Ax})_i < (\boldsymbol{Ax})_j$, 则 $f_{ij} = 0$; 如果 $(\boldsymbol{Ax})_i > (\boldsymbol{Ax})_j$, 则 $f_{ij} = 1$; 如果 $(\boldsymbol{Ax})_i = (\boldsymbol{Ax})_j$, 则 $f_{ij} = \dfrac{1}{2}$. 这会导致微分方程右端是间断的. 此时, 在由收益值 $(\boldsymbol{Ax})_i$ 的特殊排序决定的状态空间中的每个区域里, 动力学都可以简化成一个复制方程的动力学.

并不是所有的学习都是社会学习 (即学习其他人). 我们可以学习自己的经验, 例如使用迄今为止带来成功的策略. 而且, 社会学习会忽视榜样的成功. 例如, 简单地模仿那些使用频率最高的策略.

值得强调的是, 模仿 (就好像是基因选择) 不会产生任何新的东西. 若群体中不存在策略 $\boldsymbol{e}_i$, 则永远不存在 (即存在 t, 使 $x_i(t) = 0$, 那么对任意的 t, 有 $x_i(t) = 0$). 存在新型的博弈动力学. 例如, 聪明的参与者会采取收益最高的策略, 即使在群体中尚未被使用. 如果我们假设以稳定的速率随机改变策略, 另一种突变动力学将会出现. 这可以解释为 “探索率”, 它对应遗传学上的变异.

2.8　复制方程的基本性质

容易看出, 如果在所有收益项 $(\boldsymbol{Ax})_i$ 上添加任意函数 $f(\boldsymbol{x})$, 复制方程 (2.24) 仍然不会改变: 因为 $\sum x_i = 1$, 在收益上加的项, 都会出现在平均收益 $\boldsymbol{x} \cdot \boldsymbol{Ax}$ 上, 所以做差时这两项会抵消. 特别地, 这蕴含着可以添加常数 c_j 到矩阵 $\boldsymbol{A}$ 的第 j 列 $(j = 1, \cdots, n)$, 而不会改变 S_n 上的复制动力学. 我们将频繁使用这个方法来简化分析.

另一个有用的性质是商法则: 如果 $x_j > 0$, 商的时间导数满足

$$\left(\frac{x_i}{x_j}\right)^{\cdot} = \left(\frac{x_i}{x_j}\right)[(\boldsymbol{Ax})_i - (\boldsymbol{Ax})_j]. \tag{2.31}$$

更一般地, 如果 $V = \prod x_i^{p_i}$ 那么

$$\dot{V} = V\left[\boldsymbol{p} \cdot \boldsymbol{Ax} - \left(\sum p_i\right)\boldsymbol{x} \cdot \boldsymbol{Ax}\right]. \tag{2.32}$$

复制方程的驻点 $\boldsymbol{z}$ 是所有收益值 $(\boldsymbol{Az})_i$ 相等的点, 其中 i 是满足 $z_i > 0$ 的所有下标. 这些收益都相等, 且等于平均收益 $\boldsymbol{z} \cdot \boldsymbol{Az}$. 特别地, 单形 S_n 上的所有顶点 $\boldsymbol{e}_i$ 都是驻点. (显然, 如果所有的参与者具有同样的类型, 模仿不会导致改变.) 如果线性方程组

$$(\boldsymbol{Ax})_1 = \cdots = (\boldsymbol{Ax})_n \tag{2.33}$$

存在一个解 (在 $\mathrm{int}S_n$ 上), 那么复制方程在 $\mathrm{int}S_n$ 上存在一个驻点.

类似地, 每个面上的驻点都可以通过解对应的线性方程组获得. 一般地, 每一个子单形 (包括 S_n 自己) 的内部至多有一个驻点.

可以证明, 如果 S_n 的内部没有驻点, 那么当 $t \to \pm\infty$ 时, S_n 内部的轨道都趋向于边界. 特别地, 如果策略 $\boldsymbol{e}_i$ 是**严格劣势策略** (strictly dominated strategy), 即如果存在 $\boldsymbol{w} \in S_n$ 使得对任意的 $\boldsymbol{x} \in S_n$ 都有 $(\boldsymbol{Ax})_i < \boldsymbol{w} \cdot \boldsymbol{Ax}$, 那么当 $t \to +\infty$ 时, 有 $x_i(t) \to 0$. 相反地, 如果存在一个轨道从 S_n 的边界逃离 (即存在 $a > 0$, 使得对任意的 $t > 0$, $i = 1, \cdots, n$ 有不等式 $x_i(t) > a$), 那么 $\mathrm{int}S_n$ 中存在一个驻点. 注意到对 $i = 1, \cdots, n$, 有

$$(\log x_i)^{\cdot} = \dot{x}_i / x_i = (\boldsymbol{Ax}(t))_i - \boldsymbol{x}(t) \cdot \boldsymbol{Ax}(t). \tag{2.34}$$

从 0 到 T 积分, 再除以 T, 使得等式左端变成 $[\log x_i(T) - \log x_i(0)]/T$, 而当 $T \to +\infty$ 时, 它趋近于 0. 等式右端相应的极限蕴含着时间平均

$$z_i(T) = \frac{1}{T} \int_0^T x_i(t) \mathrm{d}t \tag{2.35}$$

的聚点 z_i, 使得关系 $z_i \geqslant a > 0$, $\sum z_i = 1$ 和

$$\sum a_{1j} z_j = \cdots = \sum a_{nj} z_j \tag{2.36}$$

必须成立. 因此 $\boldsymbol{z}$ 是 $\mathrm{int}S_n$ 中的驻点.

2.9 两策略情形

考虑群体中只有两种类型的复制方程. 2×2 收益矩阵每列减去对角项, 对应的方程不会改变, 因此不失一般性可以假设矩阵 $\boldsymbol{A}$ 具有如下形式

$$\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}. \tag{2.37}$$

因为 $x_2 = 1 - x_1$, 所以只需考虑 x_1, 记 x_1 为 x. 这样 $x_2 = 1 - x$, 且

$$\dot{x} = x[(\boldsymbol{Ax})_1 - \boldsymbol{x} \cdot \boldsymbol{Ax}] = x[(\boldsymbol{Ax})_1 - (x(\boldsymbol{Ax})_1 + (1 - x)(\boldsymbol{Ax})_2)], \tag{2.38}$$

从而

$$\dot{x} = x(1 - x)[(\boldsymbol{Ax})_1 - (\boldsymbol{Ax})_2]. \tag{2.39}$$

因为 $(\boldsymbol{Ax})_1 = a(1 - x)$ 且 $(\boldsymbol{Ax})_2 = bx$, (2.39) 化简为

$$\dot{x} = x(1 - x)[a - (a + b)x]. \tag{2.40}$$

注意到

$$a = \lim_{x \to 0} \frac{\dot{x}}{x}. \tag{2.41}$$

因此, a 等于趋于消失的策略 e_1 的单位增长率的极限. 或者

$$a = \frac{\mathrm{d}\dot{x}}{\mathrm{d}x}, \tag{2.42}$$

其中导数为 $x=0$ 处的取值.

我们不考虑平凡的情形 $a=b=0$: 这种情况下, 状态空间 S_2 的所有点 (即, 区间 $0 \leqslant x \leqslant 1$) 都是驻点. 微分方程的右端是三个因式的乘积, 第一、二个因式的根分别是 0, 1, 当且仅当 $ab>0$ 时, 第三个因式在开区间 $[0,1]$ 上有一个零点 $\hat{x} = \dfrac{a}{a+b}$. 这样, 我们得到三种可能的情况, 如图 2.2 所示.

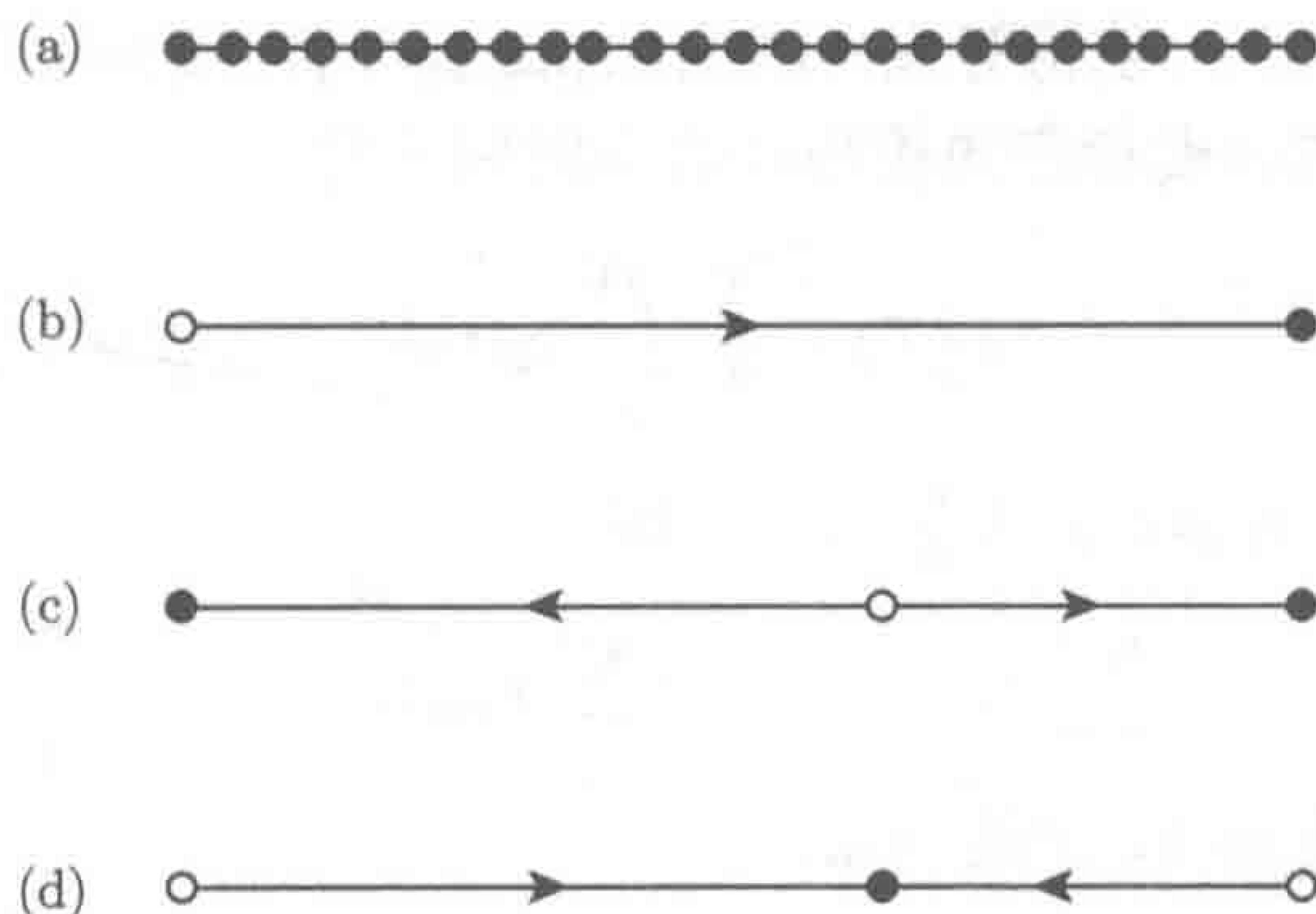


图 2.2　$n=2$ 的复制方程分类: (a) 惰性; (b) 占优; (c) 双稳态; (d) 稳定共存. 空心代表驻点. 实心代表稳定驻点

1. 当且仅当 $ab \leqslant 0$ 时, 状态空间内部没有驻点. 在这种情况下, $\dot{x}$ 在 $[0,1]$ 区间内总是有相同的符号. 如果符号是正的 (即, 如果 $a \geqslant 0$ 且 $b \leqslant 0$, 至少有一个不等号严格成立), 这意味着对每一个在 $]0,1[$ 区间上的初值 $x(0)$, 当 $t \to +\infty$ 时, 都有 $x(t) \to 1$. e_1 对 e_2 **占优** (dominate). 当 $x \in [0,1]$ 时, 它总是最佳响应策略. 反过来, 如果 $\dot{x}$ 的符号是负的, 那么 $x(t) \to 0$ 且 e_2 占优. 在这两种情况下, 占优策略都趋于驻点.

作为例子, 我们考虑 1.3 节的囚徒困境. 其收益矩阵可以转换成如下形式

$$\begin{pmatrix} 0 & -5 \\ 5 & 0 \end{pmatrix}. \tag{2.43}$$

在此情况下, 背叛占优.

2. 在 $[0,1]$ 区间上存在一个驻点 $\hat{x}$ (即 $ab>0$), 且 a 和 b 均为负. 在这种情况下, 当 $x \in [0,\hat{x}]$ 时, $\dot{x}<0$; 当 $x \in [\hat{x},1]$ 时, $\dot{x}>0$. 这意味着轨道远离 $\hat{x}$: 这个驻点

是不稳定的. 如果是前面那种情形, 一个策略将会消失: 但在**双稳态** (bistable) 情形下, 结果会依赖于初始条件. 如果 x 比临界值 $\hat{x}$ 大, 它就会增长; 如果比临界值小, 它就会变为零——这是一个正反馈.

作为例子, 我们考虑 1.5 节的重复囚徒困境. 其收益矩阵可以转换成如下形式

$$\begin{pmatrix} 0 & -5 \\ -45 & 0 \end{pmatrix}. \tag{2.44}$$

在这种情况下, 如果使用 TFT 策略的参与者超过 10%, 那么最好使用 TFT 策略.

3. 在 $[0,1]$ 区间上存在一个驻点 $\hat{x}$ (即 $ab>0$), 且 a 和 b 都为正. 在这种情况下, $x \in [0,\hat{x}]$ 时, $\dot{x}>0$; $x \in [\hat{x},1]$ 时, $\dot{x}<0$. 这个负反馈意味着当 $t \to +\infty$ 时, $x(t) \to \hat{x}$: 这个驻点 $\hat{x}$ 是一个稳定吸引子. 没有策略能消除其他策略: 它们的频度趋于**稳定共存** (stable coexistence).

下面的例子为 1.4 节的雪堆博弈. 其收益矩阵可以转变为

$$\begin{pmatrix} 0 & 10 \\ 15 & 0 \end{pmatrix}. \tag{2.45}$$

在这种情况下, 驻点对应于四成的帮助和六成的逃避.

这三种情况 (占优、双稳态和稳定共存) 将在下一节再次涉及. 现在, 我们先将复制动力系统和纳什均衡概念联系起来.

2.10 纳什均衡与饱和驻点

考虑具有对称纳什均衡 $\boldsymbol{z}$ 的 $n \times n$ 对称博弈 $(\boldsymbol{A}, \boldsymbol{A}^{\mathrm{T}})$. 即对任意的 $\boldsymbol{x} \in S_n$, 有

$$\boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{z} \leqslant \boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{z}. \tag{2.46}$$

而当 $\boldsymbol{x} = \boldsymbol{e}_i$ 时, 可得

$$(\boldsymbol{A}\boldsymbol{z})_i \leqslant \boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{z}, \ i = 1, \cdots, n. \tag{2.47}$$

如 2.2 节, 对于所有 $z_i > 0$ 的下标 i, 等号必须成立. 因此, $\boldsymbol{z}$ 是复制动力系统的一个驻点. 而且, 它还是一个**饱和驻点** (saturated rest point): 由定义, 当 $z_i = 0$ 时, 有

$$(\boldsymbol{A}\boldsymbol{z})_i - \boldsymbol{z} \cdot \boldsymbol{A}\boldsymbol{z} \leqslant 0. \tag{2.48}$$

反之, 每一个饱和驻点都是纳什均衡. 这两个概念是等价的.

马上会看到, 在 S_n 内部的每一个驻点都是饱和的; 但是在边界, 有些驻点可能不饱和. 在这种情况下, 存在策略在群体 $\boldsymbol{z}$ 中没有表现出来, 它比平均好 (事实

上, 比现存的每一种类型都好). 驻点和纳什均衡均满足: 存在 c 使得 $z_i > 0$ 时, $(\boldsymbol{Az})_i = c$; 对纳什均衡来说, 还需要额外的条件, $z_i = 0$ 时, $(\boldsymbol{Az})_i \leqslant c$.

因此, 对称纳什均衡都是驻点, 但反之不成立. 我们来看前面章节的一个例子. 显然驻点 $\hat{x} \in [0,1]$ 是纳什均衡. 情形 (1), 占优策略是纳什均衡, 其它不是. 情形 (2), 两个纯策略都是纳什均衡. 情形 (3), 纯策略都不是纳什均衡. 如果参与一个双稳态博弈, 最好选择与对手相同的策略; 但如果是一个稳定共存的博弈, 参与者应该选择与对手相反的策略. 然而, 在这两种情形两个参与者对如何选择策略可能有不同的思考.

在双稳态情形, 应该选择两个纯策略 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 中的哪一个? 首先的想法是: 选收益更高的 (如果存在的话). 这被称为**帕累托最优** (Pareto-optimal). 在 1.7 节给出的例子中, 显然 TFT 策略是帕累托最优策略. 帕累托最优的定义依赖于实际收益值, 而且无法用复制方程确定: 因为在收益矩阵的每一列加入常量后, 其它策略可能成为帕累托最优.

帕累托最优解不总是有说服力的. 例如, 考虑收益矩阵

$$\begin{pmatrix} 2 & -1000 \\ 0 & 1 \end{pmatrix}. \tag{2.49}$$

显然 $\boldsymbol{e}_1$ 是帕累托最优. 但是你能用其对抗未知对手吗? 那个参与者也许是个傻子, 选了 $\boldsymbol{e}_2$. 在这种情况下, 你会输得更多. 显然, $\boldsymbol{e}_2$ 是更安全的纳什均衡. (另一种思考是你的对手不是傻子, 只是怀疑你可能是傻子; 或者怀疑你可能怀疑, 等等.)

在一个双稳态博弈

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{2.50}$$

($\alpha > \gamma$ 且 $\delta > \beta$) 中, 如果在对手使用 $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ 策略的可能性相同的情况下, 参与者使用 $\boldsymbol{e}_1$ 提供了更高的收益, 那么 $\boldsymbol{e}_1$ 被称为**风险占优** (risk-dominant). 这意味着 $(1/2)(\alpha+\beta) > (1/2)(\gamma+\delta)$, 或

$$\gamma - \alpha < \beta - \delta. \tag{2.51}$$

此条件在每列加常量的情况下是不变的, 即在规范化矩阵 (2.37) 中有 $a > b$, 从而 (因为两个值都为负) $\hat{x} < 1/2$. 因此, 在 2×2 的双稳态博弈中, 风险占优有一个更大的吸引盆.

复制系统中极少数的关于纳什均衡和驻点的结果被认为是**进化博弈理论的民间定理** (folk theorem of evolutionary game theory). 例如, 当 $t \to +\infty$ 时, 从 S_n 内部出发的解 $\boldsymbol{x}(t)$ 的任意极限点都是纳什均衡; 并且任意稳定驻点都是纳什均衡. (如果对 $\boldsymbol{z}$ 的任意邻域 U, 存在 $\boldsymbol{z}$ 的邻域 V, 当 $\boldsymbol{x}(0) \in V$ 时, 对任意的 $t \geqslant 0$, $\boldsymbol{x}(t) \in U$

成立, 那么驻点 $\boldsymbol{z}$ 被称为是稳定的.) 这两个结果基于如下事实: 如果 $\boldsymbol{z}$ 不是纳什均衡, 则存在 i, ϵ 使得对 $\boldsymbol{z}$ 附近的 $\boldsymbol{x}$ 都有 $(\boldsymbol{Ax})_i - \boldsymbol{x}\cdot\boldsymbol{Ax} > \epsilon$. 另一方面, 如果 $\boldsymbol{z}$ 是严格纳什均衡, 那么 $\boldsymbol{z}$ 是一个渐进稳定驻点 (即不只是稳定的, 还是**吸引的** (attracting). 吸引的意思是, 对 $\boldsymbol{z}$ 的某个邻域 U, $\boldsymbol{x}(0)\in U$ 蕴含着当 $t\to+\infty$ 时, $\boldsymbol{x}(t)\to\boldsymbol{z}$). 反之, 不一定成立.

2.11 纳什均衡的存在性

为了证明 $n\times n$ 矩阵 $\boldsymbol{A}$ 的对称博弈存在对称纳什均衡, 即证明相应的复制方程 (2.24) 存在饱和驻点, 我们在方程的右端都增加一个常量 $\epsilon>0$. 当然, $\sum \dot{x}_i=0$ 不再成立. 我们通过在每一个增长率 $(\boldsymbol{Ax})_i - \boldsymbol{x}\cdot\boldsymbol{Ax}$ 上减去 $n\epsilon$ 来补偿. 因此, 我们考虑

$$\dot{x}_i = x_i[(\boldsymbol{Ax})_i - \boldsymbol{x}\cdot\boldsymbol{Ax} - n\epsilon] + \epsilon. \tag{2.52}$$

显然, $\sum \dot{x}_i = 0$ 再一次满足了. 另一方面, 若 $x_i=0$, 则 $\dot{x}_i=\epsilon>0$. 这个引入项改变了复制方程的向量场: 在 S_n 的边界 (这个边界相对未受扰动的复制方程是不变的), 受扰动方程的向量场指向内部.

我们看到 (2.52) 在 $\mathrm{int}S_n$ 中至少有一个驻点, 我们用 $\boldsymbol{z}_\epsilon$ 表示. 它满足

$$(\boldsymbol{Az}_\epsilon)_i - \boldsymbol{z}_\epsilon\cdot\boldsymbol{Az}_\epsilon = \epsilon\left(n - \frac{1}{(\boldsymbol{z}_\epsilon)_i}\right). \tag{2.53}$$

令 ϵ 趋于 0, 假设 $\boldsymbol{z}$ 是 S_n 中 $\boldsymbol{z}_\epsilon$ 的聚点. 等式左端的极限存在, 为 $(\boldsymbol{Az})_i - \boldsymbol{z}\cdot\boldsymbol{Az}$. 因此, 在 $\epsilon\to 0$ 时, 右端也应该有一个极限. 当 $z_i>0$ 时, 极限为 0; 当 $z_i=0$, 极限 $\leqslant 0$. 这意味着 $\boldsymbol{z}$ 是 (未受扰动的) 复制方程 (2.24) 的饱和驻点, 因此, 是纳什均衡.

剩下的就是证明方程 (2.52) 驻点的存在性. 知道 Brouwer 不动点定理的读者不需要再看证明. 其他读者可以参见下面两节.

2.12 斯波纳引理

考虑 $n-1$-维单形 S, 即 n 个点 $\boldsymbol{y}_1,\cdots,\boldsymbol{y}_n$ 组成的闭凸包, 其向量 $\boldsymbol{y}_i-\boldsymbol{y}_n$, $i=1,\cdots,n-1$ 是线性无关的. 向量 $\boldsymbol{y}_1,\cdots,\boldsymbol{y}_n$ 的任意非平凡子集张成 S 的**子单形**. S 的边界是 n 个全 (full) 面 ($n-2$-维) 的并. S 的**单纯分解** (simplicial decomposition) 由有限个 $n-1$-维单形组成, 这些 $n-1$-维单形的并为 S 且它们的内部不交. 进一步还需要, 如果两个 (闭的) 子单形相交, 则它们共有一个面: 即如果交集包含一个子单形的顶点, 那么这个顶点也是另一个子单形的顶点. 如图 2.3 所示.

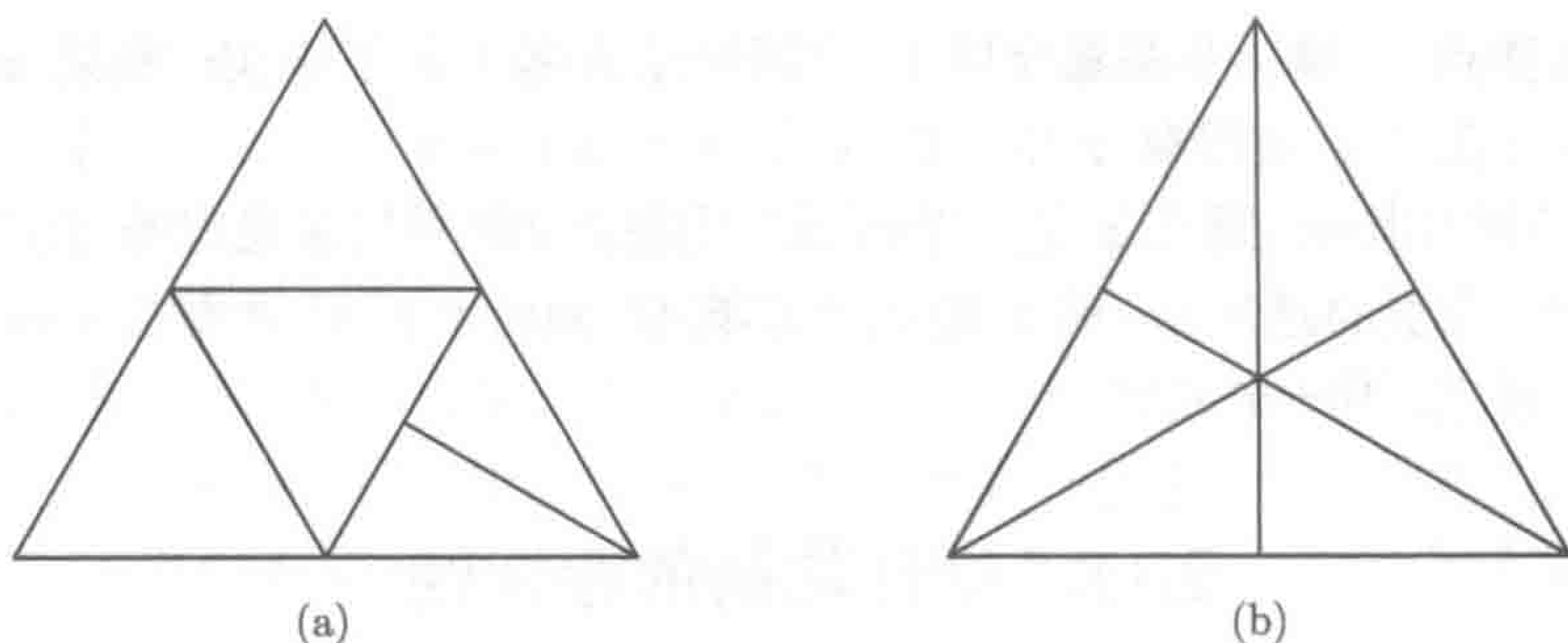


图 2.3　(a) 非单纯分解 (b) 重心分解是单纯分解

重心剖分就是一个例子. (单形 S 的**重心** (barycenter) 是 $(\boldsymbol{y}_1+\cdots+\boldsymbol{y}_n)/n$). 先考虑 1-维子单形的重心, 即边的中点. 它们将 S 的边划分为 1-维子单形. 再引入 S 的 2-维子单形的重心, 将 2-维子单形作为一个凸包考虑, 该凸包由重心和相应面边界上的 1-维子单形组成; 高维的可以继续这样考虑.

假设单形的顶点用 n 种颜色来**染色** (coloring), 意思是, 对于每一个顶点给予一个颜色 $i\in\{1,\cdots,n\}$. 我们将 S 的顶点 $\boldsymbol{y}_i$ 染成颜色 i, 且对于 S 的任意子单形, 只考虑张成子单形的顶点的那些颜色. 如果 $I\subset\{1,\cdots,n\}$ 是所有实际出现在这个子单形顶点上的颜色的列表, 我们称这个子单形为 I-染色.

斯波纳引理 (Sperner's lemma) 表明了单纯分解上总是存在奇数个 $\{1,\cdots,n\}$-染色子单形. (特别地, 只需要全染色集至少出现在 S 的一个子单形上.)

用归纳法证明. 对于 $n=2$ (即 S_2 的一个剖分) 显然成立. 假设 $n-1$ 的情况成立. 我们可以把这个情况应用到 S 上与 $\boldsymbol{y}_n$ 对立的边界面: 它的单纯分解包含奇数个 $\{1,\cdots,n-1\}$-染色子单形.

现在构造一个图, 其顶点是 S 的子单形的重心. 当且仅当子单形共有一个 $\{1,\cdots,n-1\}$-染色面时, 我们将其重心连接起来, 如图 2.4 所示. 我们在 S 的外边加一个顶点 o, 将它和那些在 S 的边界上共有一个 $\{1,\cdots,n-1\}$-染色面的子单形的重心连接起来. 我们马上看到 o 连接了奇数个重心, 这些重心所属的子单形都有一个全面, 它属于 S 且对立于 $\boldsymbol{y}_n$.

如果子单形是 $\{1,\cdots,n\}$-染色, 则它恰好有一个 $\{1,\cdots,n-1\}$-染色面. 因此, 它的重心位于图上的一条边; 它是图的端点. 对于其他重心来说, 它们要么连接两条边, 要么不连接. 确实, 如果一个不是 $\{1,\cdots,n\}$-染色的子单形有一个 $\{1,\cdots,n-1\}$-染色面, 那么剩下的顶点必须有 $1,\cdots,n-1$ 种颜色中的一种, 因此, 此子单形恰好有一个额外的 $\{1,\cdots,n-1\}$-染色面.

注意到图可能有闭环. 但是因为有奇数条边从 o 流出, 图必须有奇数个端点, 所以有奇数个 S 的全染色子单形. 因此斯波纳引理得证.

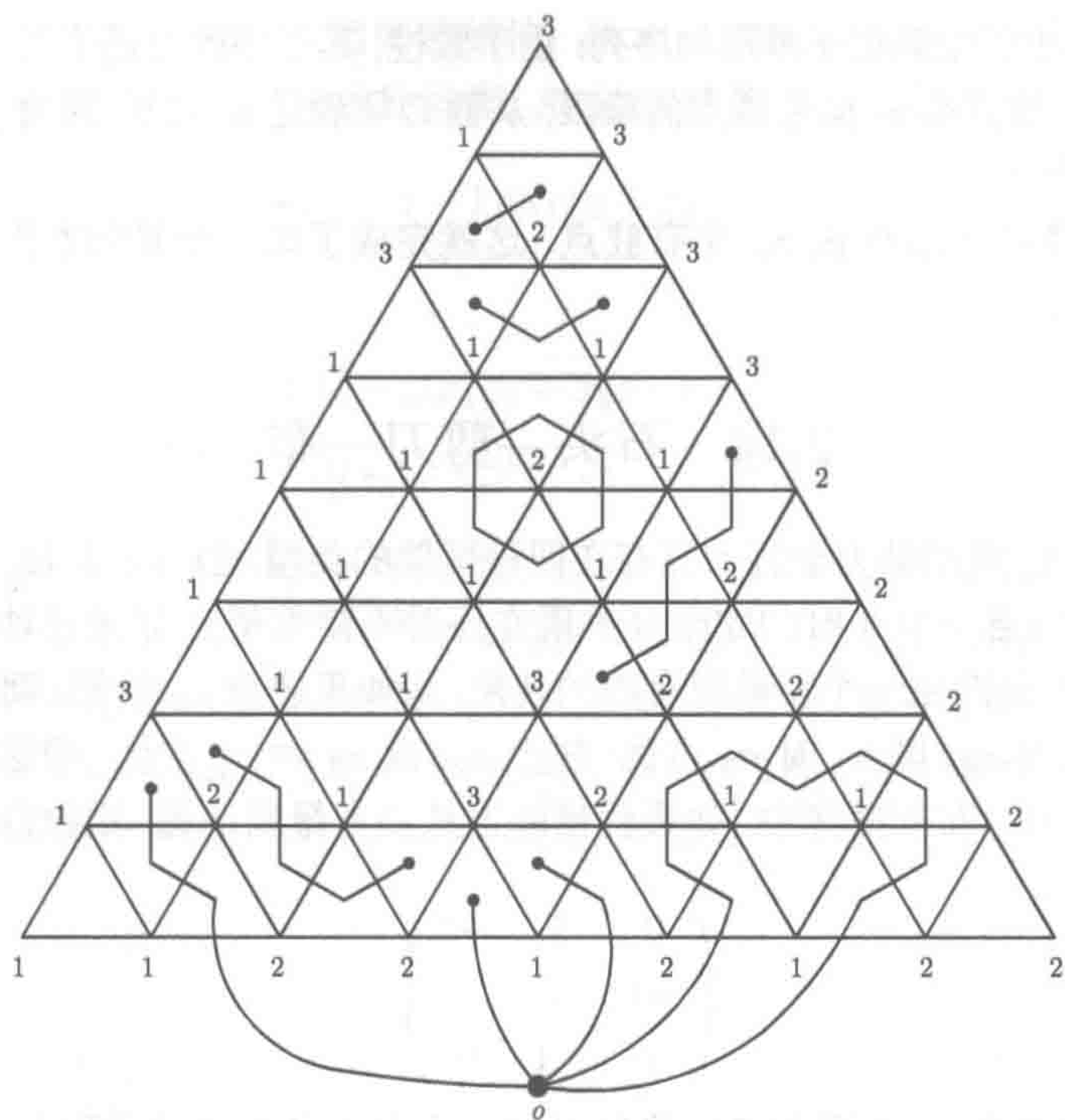


图 2.4 S_3 单纯分解的染色示意图. 描述见 2.12 节

2.13 不动点定理

现在证明在面 $\sum x_i = 1$ 上的光滑向量场如果满足

$$\sum \dot{x}_i = 0 \tag{2.54}$$

且

$$x_i = 0 \Rightarrow \dot{x}_i > 0, \tag{2.55}$$

则在 $\mathrm{int}S_n$ 中存在一个驻点. 我们利用反证法证明, 假设它没有驻点. 对每一点 $\boldsymbol{x} \in S_n$, 可以标记其“颜色”为 $i := \min\{j : \dot{x}_j < 0\}$. 因为 $\dot{\boldsymbol{x}} \neq \mathbf{0}$ 且 $\sum_j \dot{x}_j = 0$, 所以这是可能的. 注意, 条件 (2.55) 意味着 S_n 的每一个子单形, 只有张成该面的那些顶点的颜色被用到.

这引出了对任意单纯分解的一种染色法. 每一个这样染色的单纯分解都有奇数个全染色子单形. 现在考虑单纯分解序列其最大子单形的直径趋于 0. (例如, 我们可以对 S 进行重心剖分, 然后无限的迭代下去.)

这样就产生了全染色子单形的序列: 由于紧性, 其子序列会趋于某一点 $\boldsymbol{x} \in S_n$. 对于每一个 i, 这点是 i- 染色顶点的极限, 从而必须满足 $\dot{x}_i \leqslant 0$. 因为 $\sum \dot{x}_i = 0$ 意味着 $\dot{\boldsymbol{x}} = \mathbf{0}$, 矛盾.

因此, 向量场 (2.52) 在 S_n 中有驻点. 这就完成了每一个复制方程都有一个饱和驻点的证明.

2.14　石头—剪刀—布

对应 $n = 2$, 复制动力学行为只存在四种可能的类型, 当 $n = 3$ 时, 则存在上百种可能的类型 (当 $n > 3$ 时的完全分类现在还看不到希望). 如果三种策略以循环方式占优的话, 会产生一个特殊而有趣的现象. 即如果缺少 $\boldsymbol{e}_3$ 的话, 则 $\boldsymbol{e}_1$ 对 $\boldsymbol{e}_2$ 占优; 同样的, 缺少 $\boldsymbol{e}_1$, 则 $\boldsymbol{e}_2$ 对 $\boldsymbol{e}_3$ 占优; 缺少 $\boldsymbol{e}_2$, 则 $\boldsymbol{e}_3$ 对 $\boldsymbol{e}_1$ 占优. 像这样的循环出现在石头—剪刀—布的博弈中. 如果假设胜方从负方得到 1 元, 则收益矩阵可以表示为

$$\begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}. \tag{2.56}$$

这是一个零和博弈: 一个参与者得到的是另一个参与者失去的. 因此, 群体中的平均收益 $\boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x}$ 为 0. 它有四个驻点, 一个位于中心 $\boldsymbol{m} = (1/3, 1/3, 1/3) \in \mathrm{int}S_3$, 其他三个分别位于顶点 $\boldsymbol{e}_i$. $\boldsymbol{m}$ 是唯一的纳什均衡.

考虑函数 $V := x_1x_2x_3$, 在 S_3 的内部为正 (在 $\boldsymbol{m}$ 达到最大值), 在边界上为零. 利用 (2.32) 我们知道, 映射 $t \to V(\boldsymbol{x}(t))$ 满足

$$\dot{V} = V(x_2 - x_3 + x_3 - x_1 + x_1 - x_2) = 0. \tag{2.57}$$

因此, V 是一个*常数运动* (constant of motion): 所有复制方程的轨道 $t \to \boldsymbol{x}(t)$ 保持在 V 的常数水平集, 见图 2.5. 这意味着 S_n 内部的所有轨道都是围绕 $\boldsymbol{m}$ 的闭轨. 由 S_3 的三个顶点 $\boldsymbol{e}_i$ 和沿着 S_3 的边连接三个顶点的轨道组成的不变集形成了一个*异宿环* (heteroclinic set). 异宿环上的任意两个点可以通过 "跟踪动力学" 的方法连接起来. 即, 只需沿着异宿环轨道移动且在极少的合适的时机迈出任意一小步. 此时, 它意味着沿着一条朝向 $\boldsymbol{e}_1$ 的边移动, 然后走到离开 $\boldsymbol{e}_1$ 的边. 这一步可以任意小: 只需要等到足够接近 "连接点" $\boldsymbol{e}_1$.

现在考虑具有如下矩阵的博弈

$$\begin{pmatrix} 0 & a & -1 \\ -1 & 0 & a \\ a & -1 & 0 \end{pmatrix}. \tag{2.58}$$

当 $a>0$ 时, 它跟循环占优有相同的结构. 当 $a\neq 1$ 时, 博弈不再是零和博弈, 但是它有同样的驻点. 如前, $\boldsymbol{m}$ 是纳什均衡, S_3 的边界是一个异宿环. 此时,

$$\boldsymbol{x}\cdot \boldsymbol{A}\boldsymbol{x}=(a-1)(x_1x_2+x_2x_3+x_3x_1), \tag{2.59}$$

因此

$$\dot{V}=V(a-1)[1-3(x_1x_2+x_2x_3+x_3x_1)] \tag{2.60}$$

$$=\frac{V(a-1)}{2}[(x_1-x_2)^2+(x_2-x_3)^2+(x_3-x_1)^2]. \tag{2.61}$$

这个表达式在 S_3 的边界和 $\boldsymbol{m}$ 点为零, 而在 S_3 的其他地方与 $a-1$ 同号. 如果 $a>1$, 这意味着所有穿过 V 的等位线的轨道都是向上的, 因此趋向于 $\boldsymbol{m}$. 这蕴含着群体中三种类型最终都会以相同的频度存在: 驻点 $\boldsymbol{m}$ 渐近稳定, 如图 2.5(b) 所示. 但当 $a<1$ 时, 轨道都是向下的, 趋向于 S_3 的边界. 纳什均衡 $\boldsymbol{m}$ 为不稳定驻点, 边界上的异宿环吸引了所有的其他轨道, 如图 2.5(c) 所示.

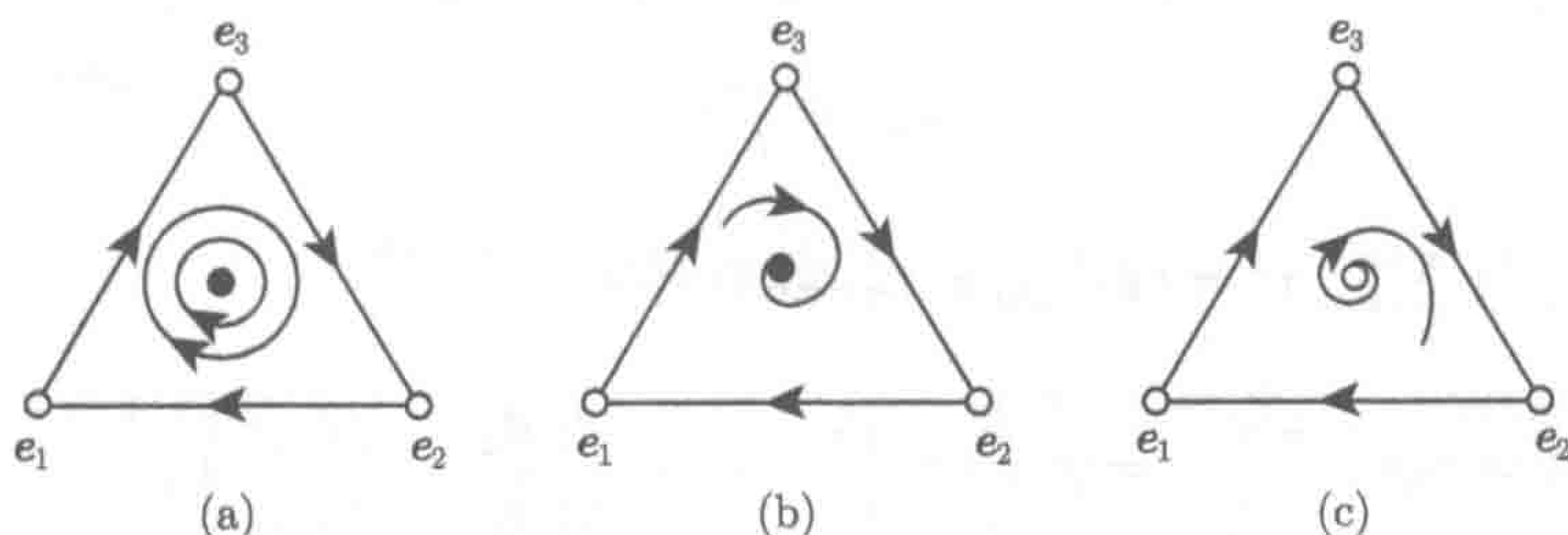


图 2.5 收益矩阵为 (2.58) 的石头—剪刀—布博弈的复制动力系统: (a) $a=1$; (b) $a>1$; (c) $0<a<1$

现在考虑当 $a<1$ 时的群体状态 $\boldsymbol{x}(t)$. 如果状态非常接近顶点, 例如 $\boldsymbol{e}_1$, 即接近驻点, 因此几乎是静止的. 在很长一段时间, 这种状态都不会明显改变. 然后, 它加快速度, 朝着顶点 $\boldsymbol{e}_3$ 移动, 临近时它会减速, 然后这种状态持续很长一段时间, 等等. 这就像一个 "间断均衡" 的周期形式: 长期的准静止状态随后出现突然的剧变.

2.15 随机过程与固定概率

迄今为止, 我们都是考虑无限大群体的极限情况. 如果考虑一个有限大的群体, 就不能再利用确定性的模型. 在有限群体中, 应该考虑由于取样所产生的随机变动. 我们需要使用随机过程来代替常微分方程.

考虑最简单的情形, 假设一个大小为 M 的有限群体由两种类型的参与者 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 组成. 通常, 其中一个参与者通过模仿群体中选定的类型来更新策略. 群体的

状态是由类型 e_1 个体数量 i 确定的 (当然类型 e_2 的参与者数量就为 $M-i$). 令 p_{ij} 为从状态 i 到状态 j 的转移概率. 矩阵 P 是一个三对角阵, 即如果 $|j-i|>1$, 那么 $p_{ij}=0$. 状态 0 和 M 是吸收的: 如果所有的个体具有相同的类型, 则模仿将不能引出其他类型. 我们令 $p_{i,i+1}=b_i$, $p_{i,i-1}=d_i$ (因为这些转移概率在另一种解释中对应**出生率** (birth rate) 和**死亡率** (death rate)).

我们用 p_i 表示群体从状态 i 最终达到全部由 e_1 类型组成的状态 M 的概率. 因为一旦达到就不会离开, 所以状态 M 是**吸收的** (absorbing). 令 p_i 是起始于状态 i, 到类型 e_1 **固定** (fixation) 发生的概率, 那么 p_i 必须满足

$$p_i = d_i p_{i-1} + (1-b_i-d_i)p_i + b_i p_{i+1}, \tag{2.62}$$

其中 $i=1,\cdots,M-1$. 实际上, 在第一次更新后, e_1 类型的数量会增加或减少一个, 或者仍然不变 (当一个参与者模仿自己的类型); 且第一步以后, 固定必须发生. 而且, 有 $p_0=0$, $p_M=1$. 令 $y_i:=p_i-p_{i-1}$, 方程 (2.62) 可以写成如下形式

$$y_{i+1} = \frac{d_i}{b_i} y_i. \tag{2.63}$$

因为 $p_1=y_1$ 且 $\sum_{i=0}^{k} y_i = p_k - p_0 = p_k$, 我们得到

$$1 = p_M = \sum_{i=1}^{M} y_i = p_1\left(1+\frac{d_1}{b_1}+\cdots+\frac{d_1}{b_1}\frac{d_2}{b_2}\cdots\frac{d_{M-1}}{b_{M-1}}\right), \tag{2.64}$$

所以

$$p_i = \frac{1+\sum_{j=1}^{i-1}\prod_{k=1}^{j} d_k/b_k}{1+\sum_{j=1}^{M-1}\prod_{k=1}^{j} d_k/b_k}, \tag{2.65}$$

其中 $i=1,\cdots,M$. 特别地, 我们用 $\rho_{1,2}$ 表示**固定概率** (fixation probability) p_1, 即在只有一个为 e_1 类型, 其余全为 e_2 类型组成的群体中, e_1 类型最终会被所有参与者模仿的概率. 它由下式给出

$$\rho_{1,2} = \frac{1}{1+\sum_{j=1}^{M-1}\prod_{k=1}^{j} d_k/b_k}. \tag{2.66}$$

到目前为止, 我们并没有详细说明模仿机制. 在本章中, 我们将考虑所谓的 Moran 过程, 它是在群体遗传学的背景下发展的. 因此, 我们将采用对应的术语, 并假设每个个体有一个确定的 "适合度". 此时, 适合度是衡量成功的尺度. 这样, 具有更高适合度的个体更有可能被模仿 (2.7 节). Moran 过程是这样的, 随机选择一个个体 (每个个体被选中的概率都为 $1/M$) 且赋予它 "模型参与者" 的类型, 从群体中选出这个类型的概率正比于这个类型的成功率.

这样, 作为第一个例子, 假设 e_1 类型个体的适合度为 r, 而 e_2 类型个体的适合度规范化为 1. 然后我们得到了死亡率

$$d_i = \left(\frac{i}{M}\right)\left(\frac{M-i}{ri+M-i}\right), \tag{2.67}$$

其中, 第一个分式是被更新的个体为 e_1 类型概率, 第二个分式是选择 e_2 类型的概率. 相似的, 有生存率,

$$b_i = \left(\frac{M-i}{M}\right)\left(\frac{ri}{ri+M-i}\right). \tag{2.68}$$

因此, $d_i/b_i = 1/r$, 并且

$$\rho_{1,2} = \frac{1-r^{-1}}{1-r^{-M}}. \tag{2.69}$$

令 $r \to 1$, 得到 $\rho_{1,2}$ 的极限值 $\rho_{1,2} = 1/M$. 这是**中性类型** (neutral type) 的固定概率, 即如果与原住者 e_2 个体所做的一样好, 单一 e_1 类型个体会被模仿且最终充满整个群体的概率.

2.16 有限群体中的博弈

现在假设群体的大小是 M, 个体参与成对博弈, 策略都由 e_1 或者 e_2 类型决定. 如果收益矩阵是

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \tag{2.70}$$

那么期望收益依赖于群体的状态 i. 对于 e_1 类型参与者, 由下式给出

$$F_i = \alpha\frac{i-1}{M-1} + \beta\frac{M-i}{M-1}, \tag{2.71}$$

对 e_2 类型参与者, 则为

$$G_i = \gamma\frac{i}{M-1} + \delta\frac{M-i-1}{M-1}. \tag{2.72}$$

(参与者不与自己博弈.) 在 2.7 节, 我们假设适合度, 即被模仿的可能性, 是由收益和 "基线适合度" 的凸组合给出的, 和以前一样, 把它们标准化为 1. 因此, 在状态为 i 的群体中, e_1 个体的适合度为

$$f_i = (1-s) + sF_i, \tag{2.73}$$

而 e_2 类型个体的适合度为

$$g_i = (1-s) + sG_i. \tag{2.74}$$

这里, 参数 $s \in [0,1]$ 度量了 "选择的强度", 即博弈对于整体成功的重要性. 如果 $s=0$, 适合度与博弈无关. 在无穷大群体的极限情况, Moran 过程通过由方程 (2.29) 导出了转变率, 且因此导出了复制方程.

出生率和死亡率分别为

$$b_i = \left(\frac{M-i}{M}\right)\left(\frac{if_i}{if_i+(M-i)g_i}\right) \tag{2.75}$$

及

$$d_i = \left(\frac{i}{M}\right)\left(\frac{(M-i)g_i}{if_i+(M-i)g_i}\right), \tag{2.76}$$

所以

$$\frac{d_i}{b_i} = \frac{g_i}{f_i} = \frac{1-s(1-G_i)}{1-s(1-F_i)}. \tag{2.77}$$

因此, 固定概率 (2.66) 由下式给出

$$\rho_{1,2} = \left[1+\sum_{j=1}^{M-1}\prod_{i=1}^{j}\frac{1-s+sG_i}{1-s+sF_i}\right]^{-1}. \tag{2.78}$$

2.17 极限情况

对于小的 s, 表达式 (2.77) 可以展至一阶近似

$$\frac{d_i}{b_i} = 1-s(F_i-G_i). \tag{2.79}$$

由方程 (2.71) 和 (2.72), 有

$$H_i := F_i - G_i = \frac{1}{M-1}[\bar{e}+\bar{f}i], \tag{2.80}$$

其中, $\bar{e} = -\alpha+\beta M-\delta M+\delta$, $\bar{f} = \alpha-\beta-\gamma+\delta$. 取方程 (2.78) 的一阶近似, e_1 类型的固定概率为

$$\rho_{1,2} = \left[1+\sum_{k=1}^{M-1}\prod_{i=1}^{k}(1-sH_i)\right]^{-1}. \tag{2.81}$$

容易知道

$$\sum_{k=1}^{M-1}\prod_{i=1}^{k}(1-sH_i) = M-1-s\sum_{i=1}^{M-1}(M-i)H_i, \tag{2.82}$$

且

$$\sum_{i=1}^{M-1}(M-i)(\bar{e}+\bar{f}i) = M(M-1)\bar{e}+(M\bar{f}-\bar{e})\sum_{i=1}^{M-1}i-\bar{f}\sum_{1}^{M-1}i^2. \tag{2.83}$$

右端第一项和为 $M(M-1)/2$ 而第二项和为 $M(M-1)(2M-1)/6$. 合并后有

$$M(M-1)(M\bar{f}+\bar{f}+3\bar{e})/6=M(M-1)(eM-f)/6, \tag{2.84}$$

其中 $e=\alpha+2\beta-\gamma-2\delta$, $f=2\alpha+\beta+\gamma-4\delta$. 取方程 (2.81) 的一阶近似

$$\rho_{1,2}=\left[1-\frac{s}{6}(eM-f)\right]^{-1}/M. \tag{2.85}$$

如果它的固定概率比中性漂变更高, 即 $\rho_{1,2}>1/M$, 也就是 $eM>f$, 即

$$\alpha(M-2)+\beta(2M-1)>\gamma(M+1)+\delta(2M-4). \tag{2.86}$$

则称策略 e_1 是有利 (advantageous) 的. 当 $M\to\infty$ 时, 有

$$\alpha+2\beta>\gamma+2\delta, \tag{2.87}$$

或者利用矩阵 (2.37) 的标准形式, 有 $b<2a$. 如果 e_1 关于 e_2 占优, 即 $b\leqslant 0$ 且 $a\geqslant 0$ (其中一个不等号严格成立), 则此不等式总成立. 占优策略总是有利的. 在稳定共存情形, 即 $a>0$ 且 $b>0$ 时, 它意味着 $\hat{x}>1/3$, 其中 $\hat{x}=\dfrac{a}{a+b}$ 为纳什均衡. 因此, 如果 $1/3<\hat{x}<2/3$, 则两个策略都是有利的. 最后, 在双稳态博弈中, 即如果 $a<0$ 且 $b<0$, 则不等式 (2.86) 意味着 $\hat{x}<1/3$, 其中 $\hat{x}$ 是 $]0,1[$ 中不稳定的纳什均衡. 这意味着对复制方程来说, e_1 的吸引盆比 e_2 的两倍还要大. 特别的, 如果 e_1 是有利的, 那么它是风险占优的. 如果 $\hat{x}$ 位于 1/3 和 2/3 之间, 两个策略没有一个是有利的.

对于重复的囚徒困境 (1.7 节) 或雪堆博弈 (1.4 节), 我们从不等式 (2.86) 可以看出, 两博弈分别对应 $M>4$ 和 $M>20$ 时, 合作是有利的.

e_2 的固定概率 $\rho_{2,1}$ 的获得方式和 e_1 的是类似的 (用 $-2\alpha-\beta+2\gamma+\delta$ 代替 e 并用 $-4\alpha+\beta+\gamma+2\delta$ 代替 f). 条件 $\rho_{1,2}>\rho_{2,1}$ 意味着

$$(M-2)(\alpha-\delta)>M(\gamma-\beta). \tag{2.88}$$

在 M 的极限情况, 条件可以减弱为 $\alpha-\delta>\gamma-\beta$. 这正好是 $a>b$, 即 e_1 是风险占优的.

对于任意 $\pi\in[0,1]$, 向量 $(\pi,0,\cdots,1-\pi)\in S_m$ 是模仿过程的一个平稳分布. 假设有某个概率 $\mu>0$, 参与者可以不模仿他人, 而是通过随机试验改变策略. 此时, 马尔可夫链会出现. 它描述了突变和模仿之间的相互作用. 假设 μ 足够小, 使得我们能分离两个过程的时间尺度. 这意味着大部分时间群体都处于同质状态 0 或者 M. 偶尔, 一个单独个体会尝试其他策略. 这样, 就会开始一个新的模仿过程, 导致新类型消亡或者处于它的固定状态. 在这种 “不可转移的” (adiabatic) 情况下, 其

过程的结果可以通过一个两状态 1 和 2 (这对应由 e_1 或 e_2 组成的同质群体.) 的马尔可夫链逼近, 这个马尔可夫链由如下矩阵给出

$$\begin{pmatrix} 1-\rho_{2,1} & \rho_{2,1} \\ \rho_{1,2} & 1-\rho_{1,2} \end{pmatrix}, \tag{2.89}$$

它有唯一的平稳分布, 即左特征向量

$$\left(\frac{\rho_{1,2}}{\rho_{1,2}+\rho_{2,1}}, \frac{\rho_{2,1}}{\rho_{1,2}+\rho_{2,1}}\right). \tag{2.90}$$

它描述了在大时间跨度上两种类型的流行. 特别地, 在双稳态和群体大小为 M 的情况下, 策略 e_1 是风险占优当且仅当随机过程花了更多的时间在对应的同质状态.

这种 "不可转移的" 情形在 n 类型 e_i 上仍然是成立的. 如果 "突变速率" μ 足够小, 群体总是由一个或至多两个类型组成. 如果在同质群体中, 单个个体改变成另一种类型, 那么模仿过程会使类型在下一个突变出现之前固定或消亡. 如果假设这些突变是随机搜索, 即每一个非原驻类型都有 $1/(n-1)$ 的机会出现, 就得到一个 $n\times n$ 马尔可夫链 P, 里面的元素为转移概率 p_{ij}, 且

$$p_{ij} = \rho_{j,i}/(n-1), \tag{2.91}$$

其中 $i \neq j$. 这里, $\rho_{j,i}$ 是 j 在 i 中的固定概率. 即单个 j 类型个体被模仿最终充满整个群体的概率, 且原群体只有一个 j 类型, 其余全为 i 类型.

2.18 参考文献

现在已经成为经典的进化博弈理论的第一本著作出自 Maynard Smith (1982). 以后的著作中, Weibull (1995), Hofbauer 和 Sigmund (1998), 以及 Cressman (2003) 的研究方法本质上最接近本书, 也可以参看 Colman (1995), Fudenberg 和 Levine (1998), Gintis (2000), 以及 Vincent 和 Brown (2005). Hofbauer 和 Sigmund (2003) 以及 Sandholm (2010[①]) 考虑了更一般的博弈动力学. 复制方程由 Taylor 和 Jonker (1978) 引入, 也可以参看 Hofbauer, Schuster 和 Sigmund (1979) 以及 Zeeman (1980); 复制方程的名字由 Schuster 和 Sigmund (1983) 给出. 通过社会学习的推导由 Helbing (1992) 和 Schlag (1998) 提出. Bomze (1983) 给出了三策略复制动力学行为的分类. Hofbauer (2000) 给出了纳什均衡对的概念, 存在性证明和各种博弈动力学之间的关系. 处理有限群体博弈的方法取自 Nowak (2006a), 也可以参看 Nowak 等 (2004) 和 Taylor 等 (2004). 另外一个不太 "随机" 的相关方法可参看 Kandori, Mailath 和 Rob (1993) 以及 Kaniovski 和 Young[②] (1995).

①译注: 原文此处为 "2009".
②译注: 原文此处为 "Peyton Young 和 Foster".

第 3 章 直接回报: 重复的作用

3.1 帮 助

正如达尔文写到, "人类体力的弱小、速度的缓慢和天然武器的缺乏等等缺陷, 已被其社会属性完全弥补 $\cdots$ 这些社会属性, 使得人们可以向他人提供帮助, 或从其他人处获得帮助". 其最简单的形式, 帮助意味着自己支付成本 c 使得另一个个体获得收益 b. 这可以看成是社会相互作用的基本原子.

在*捐赠博弈*(donation game) 中, 两个参与者必须同时决定 (更准确地, 在不知道对手的决定的情况下) 是否给予对手帮助. 这两个策略 e_1 和 e_2 将分别用 C (合作) 和 D (背叛) 来表示. 这样就有如下的收益矩阵:

$$\left(\begin{array}{c|c} b-c & -c \\ b & 0 \end{array}\right). \tag{3.1}$$

如果没有特别申明, 我们总假设 $b > c > 0$. 第二个策略 D 优于第一个策略. 如 1.3 节所述, 这是一个囚徒困境的例子, 一个 2×2 的对称博弈, 其收益矩阵

$$\left(\begin{array}{c|c} R & S \\ T & P \end{array}\right) \tag{3.2}$$

满足

$$T > R > P > S. \tag{3.3}$$

囚徒困境刻画了共同利益 (R 大于 P) 和私利 (D 优于 C) 之间的冲突. 利己将会在这场冲突中取胜. 事实上, 这个博弈有唯一的纳什均衡, 即背叛; 对成功者的模仿导致合作无情地终止, 见 2.10 节.

比较捐赠博弈和雪堆博弈是有趣的 (1.4 节). 如果两个参与者共同支付成本 c, 它们都能得到收益 b ($c < b$). 两个参与者必须同时决定是否支付, 并且知道如果都支付, 它们将共同承担成本. 其收益矩阵是

$$\left(\begin{array}{c|c} b-\frac{c}{2} & b-c \\ b & 0 \end{array}\right). \tag{3.4}$$

显然, 选择与对手相反的做法是最好的. 如果对手愿意支付, 你可以安全地拒绝支付. 但如果对手不愿意支付, 那么你最好支付. 显然, 参与者都不希望支付. 这个博

弈有唯一的对称纳什均衡. 它表明支付的概率是 $\frac{2(b-c)}{2b-c}=1-\frac{c}{2b-c}$. 我们注意到, 如果捐赠博弈进行两次, 那么两个参与者每次都合作的收益是轮流合作的两倍. 在两回合的雪堆博弈中, 两个参与者每次都合作与轮流合作收益一样. 雪堆博弈是懦夫博弈的一个例子, 懦夫博弈是一个 2×2 的对称博弈, 其收益矩阵 (3.2) 满足

$$T>R>S>P. \tag{3.5}$$

排序的微小差异 (S 和 P 互换) 将产生相当大的影响.

3.2 重复博弈

现在考虑多个回合的同步捐赠博弈. 如果两个参与者都知道博弈的回合数, 那么如 1.5 节所看到的, 由逆向递推预知, 利己的参与者在每一回合都应该选择背叛.

假设两个参与者不知道博弈会持续多少回合. 通常, 可以假设每一回合之后, 下一回合发生的概率为常数 $w<1$ (或者也可以假设回合数服从泊松分布). 把初始回合记为 0, 第 n 次重复的回合记为 n. 那么博弈至少重复 n 次的概率为 w^n. 博弈恰好进行 $n+1$ 回合 (初始回合后重复 n 次) 的概率为 $w^n(1-w)$. 回合数是一个服从几何分布的随机变量, 其期望值为

$$1(1-w)+2w(1-w)+\cdots+nw^{n-1}(1-w)+\cdots=\frac{1}{1-w}. \tag{3.6}$$

记第 n 回合的收益为 $A(n)$, 则总收益的期望值为

$$\sum_{n=0}^{+\infty}w^n(1-w)[A(0)+\cdots+A(n)], \tag{3.7}$$

根据 Abel 求和公式 (Abel's summation formula), 上式为幂级数 $A(0)+wA(1)+\cdots$. 由于 $A(n)\in\{R,S,T,P\}$, 所有的 $A(n)$ 一致有界, 因此表达式 (3.7) 总是收敛到某个 $A(w)$, 其中 $0\leqslant w<1$. 每一回合的平均收益为

$$(1-w)A(w)=(1-w)\sum_{n=0}^{+\infty}w^nA(n). \tag{3.8}$$

通常分析极限情形 $w=1$ 是有益的. 在这种情况下, 博弈有无穷多回合, 总收益 $A(0)+A(1)+\cdots$ 可能会发散. 此时, 考虑每一回合收益关于时间的平均就很方便, 即:

$$\lim_{n\to+\infty}\frac{A(0)+\cdots+A(n)}{n+1}, \tag{3.9}$$

只需要这个极限存在即可. 弗罗贝尼乌斯 (Frobenius) 定理表明, 在这种情形下, 表达式 (3.9) 由 (3.8) 的极限给出, 即 $\lim_{w\to1}(1-w)A(w)$.

3.3 合作者、背叛者以及回报者

首先, 考虑只有三个策略的相互作用. 合作者总是提供帮助; 背叛者总是不提供帮助; 回报者当且仅当对手前一回合不提供帮助时, 本回合不提供帮助 (默认回报者初始回合提供帮助). 这分别对应三个策略 $\boldsymbol{e}_1 = \text{AllC}$, $\boldsymbol{e}_2 = \text{AllD}$ 和 $\boldsymbol{e}_3 = \text{TFT}$.

考虑一个大的、均匀混合的群体. 三个策略的频度分别为 x, y, z $(x+y+z=1)$. 用 P_x, P_y, P_z 表示参与者使用这些策略所得总收益的期望值 (而不是前一章的 $(\boldsymbol{Ax})_1$ 等等). 该群体的平均收益为 $\bar{P} = xP_x + yP_y + zP_z$. 如 2.7 节, 假设越成功的策略越可能被模仿. 因此, 群体中这三个策略频度的演化可以由如下复制方程给出:

$$\begin{aligned} \dot{x} &= x(P_x - \bar{P}) \\ \dot{y} &= y(P_y - \bar{P}) \\ \dot{z} &= z(P_z - \bar{P}). \end{aligned} \tag{3.10}$$

我们将用到以下事实: 在每一个收益项上加上相同的函数, 复制方程 (在单形 S_3 上) 保持不变 (2.8 节), 相应的项仍用 P_x, P_y, P_z 及 $\bar{P}$ 表示. 特别地, 可以在收益矩阵的每列加上适当的常数使其标准化.

AllD 对 AllD 每一回合获得收益 $A(n) = 0$. 因此, $A(w) = 0$. TFT 对 AllD 首回合收益为 $A(0) = -c$. 此后, 对 $n \geqslant 1$, 有 $A(n) = 0$. 从而 $A(w) = -c$, 等等.

三个策略 AllC, AllD, TFT 相互作用的收益矩阵为

$$\begin{pmatrix} b-c & -c & b-c \\ b & 0 & b(1-w) \\ b-c & -c(1-w) & b-c \end{pmatrix}, \tag{3.11}$$

其中我们去掉了因子 $(1-w)^{-1}$ (即考虑每一回合的平均收益). 令 $w=1$ 则可得无限重复的情形.

在一般的囚徒困境中, 对应矩阵 (3.11) 的收益矩阵为

$$\begin{pmatrix} R & S & R \\ T & P & (1-w)T + wP \\ R & (1-w)S + wP & R \end{pmatrix}. \tag{3.12}$$

3.4 皮洛斯式胜利

仍然考虑捐赠博弈, 并且标准化相应的复制方程, 使得背叛者的收益 P_y 为 0. 这样, 有

$$P_x = -c + wbz, \quad P_z = P_x + wcy. \tag{3.13}$$

此时, $P_z - \bar{P} = yg$, 其中

$$g = w(b-c)z - c(1-w). \tag{3.14}$$

在状态空间即单形 S_3 的边界 $z=0$ (无报答者), AllD 明显占优. 在边界 $x=0$, 即由背叛者和 TFT 参与者组成的群体, 有双稳态现象. 不稳定的平衡点为 $F_{yz} = (0, 1-\hat{z}, \hat{z})$, 其中

$$\hat{z} = \frac{(1-w)c}{w(b-c)}, \tag{3.15}$$

这里 $\hat{z} < 1$, 即 $w > c/b$. 特别地, 如果

$$w > \frac{2c}{b+c}, \tag{3.16}$$

则 TFT 为风险占优 (2.10 节). 如果

$$w > \frac{3c}{b+2c}, \tag{3.17}$$

则 TFT 为有利的 (2.17 节).

如果 w 接近 1, 则 $\hat{z}$ 很小. 因此, 如果 "未来的阴影" w 足够大, 一小群 TFT 就能侵入背叛者群体.

边界 $y=0$ 由驻点构成. 显然, 由 AllC 和 TFT 参与者构成的群体会一直合作, 这两个策略互不占优. 在边界 $y=0$ 上, 满足 $z \geqslant c/wb$ 的点是纳什均衡, 其他的则不是. 对此, 只需看 $P_y - \bar{P}$, 即 $P_x = -c + wbz$ 的符号, 再由 2.10 节, 纳什均衡正好是那些饱和驻点 (即如果 $y=0$, 则 $P_y \leqslant \bar{P}$).

这个博弈的纳什均衡还包括顶点 $y=1$ (只有背叛者) 和 F_{yz}. 因为 $y>0$ 时, $P_z > P_x$. 因此, 在单形 S_3 的内部没有驻点. 容易知道, 函数

$$V = x^{\frac{1-w}{w}} z^{-\frac{1}{w}} g \tag{3.18}$$

关于系统不变, 即满足 $\dot{V} = 0$.

在 $c < wb$ 的情形, 出现了一个有趣的动力学现象, 如图 3.1 所示. $g=0$ 的部分是一条平行于边界 $z=0$ 的轨线, 其收敛到鞍点 F_{yz}, 并且将单形分成两部分. 在 $g=0$ 的下方, z 减小, y 收敛到 1, 即背叛者胜出. 在线 $g=0$ 的上方, z 增加, y 收敛到 0, 即背叛者消失.

在没有背叛者时, TFT 参与者和 AllC 参与者的任意混合都对应一个驻点, 这样的混和可以看做是有条件的利他主义者和无条件的利他主义者的混合. 假设一个偶然的小随机冲击对系统产生扰动, 这会使系统在无背叛者边界 $y=0$ 来回. 如果在 $z > c/wb$ 时随机扰动引进少量的背叛者, 背叛者立即就会被消除. 如果当

$z < (1-w)c/w(b-c)$ 时, 背叛者被引入, 它们就会占领这个区域. 但如果背叛者在"中间区域", 即

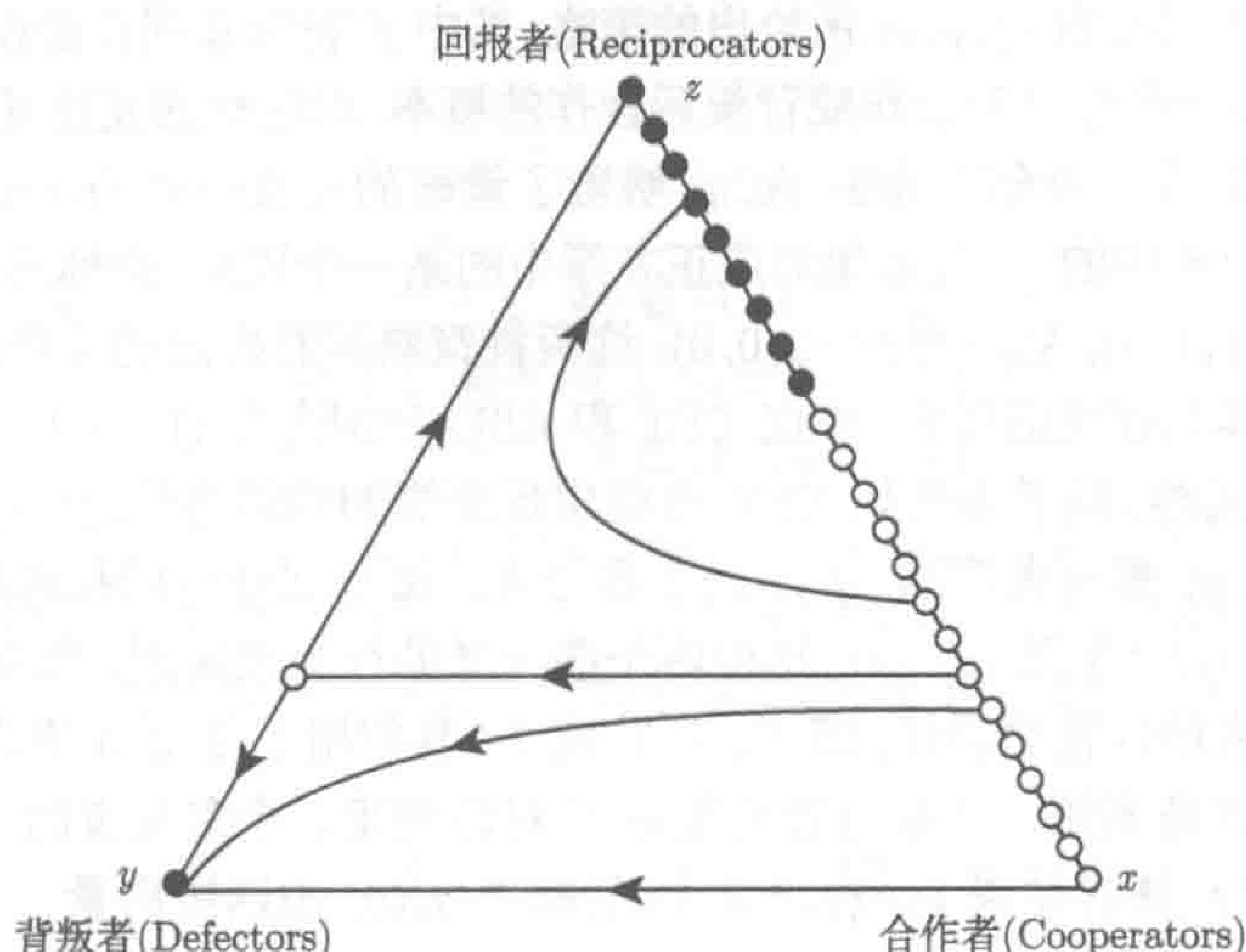


图 3.1 没有失误的合作者、背叛者和回报者. 水平线 $z = c/wb$ 将状态空间分割. 在水平线之下, 背叛者胜利; 在水平线之上, 背叛者被淘汰. 在这里及其他图形中, 实心圈对应稳定驻点, 空心圈对应不稳定驻点

$$c/wb > z > (1-w)c/w(b-c) \tag{3.19}$$

时被引入, 则背叛者先增加, 然后消亡. 在侵入阶段, AllD 参与者将会剥削并最终耗尽 AllC 参与者. 这是一种皮洛斯式的胜利: 背叛者最终将会遇到 TFT, 这是它们毁灭的原因.

从背叛者的角度来看, 当 $z > \hat{z}$ 时, 任何侵入性尝试注定是要失败的, 并且以 $y = 0$ 及 $z > c/wb$ 的状态终止. 形象地说, 背叛者唯一的希望是带着侵入的企图, 等待一系列小的随机冲击, 这种漂变将群体沿着边界 $y = 0$ 移动到 $z < \hat{z}$ 的区域. 漂变需要时间. 如果侵入企图出现得太频繁, 漂变永远没有足够的时间将群体引入到有利于背叛者的区域. 因此, 背叛者不应该太频繁地企图侵入. 换句话说, 只有当背叛者的侵入企图足够频繁时合作者才是安全的. 如果侵入企图太少, 合作型社会会失去其免疫力——随机波动使得群体中回报者的数量太少而不能击退少数背叛者的侵入.

3.5 反 应 策 略

到目前为止, 我们总是假定参与者无误地实施其意图. 如果假定参与者偶尔出

错, 则会得到非常不同的结果. 这引出了对随机策略的研究, 随机策略由每一回合合作与否的概率来描述.

首先考虑由三元组 (f,p,q) 所给出的策略, 其中 f 表示 0 回合合作的概率, p 与 q 分别表示在前一回合对手合作或背叛后合作的概率. 对这样的*反应策略*, 合作倾向完全取决于对手前一回合的做法. (p,q) 明确了策略的*反应规则* (reaction norm) 是单位正方形 $[0,1]^2$ 中的一点, 如果对应正方形中的某一个顶点, 则称其是*确定的*. 比如 TFT 对应 $(1,1,0)$, AllD 对应 $(0,0,0)$; 这两种策略都有确定的反应规则. 以 (小) 概率 ϵ 实施与本意相反的行为, 将使 TFT 和 AllD 分别变为 $(1-\epsilon,1-\epsilon,\epsilon)$、$(\epsilon,\epsilon,\epsilon)$. 记 $\rho:=p-q$. 显然, 除了某些如 TFT 有确定反应规则的策略外, $|\rho|<1$.

考虑 (f,p,q) 参与者遭遇 (f',p',q') 参与者. 每一回合有四种可能的结果, 即 (C,C), (C,D), (D,C) 及 (D,D), 这由两个参与者的行为来决定. 博弈结果也可以用第一个参与者的收益来描述, 即 R, S, T 或 P, 我们用 1, 2, 3, 4 来表示收益结果所在的状态. (注意到第一个参与者的收益 S 对应到第二个参与者的 T.)

在初始回合, 博弈结果 $i\in\{1,2,3,4\}$ 的概率 $x_i(0)$ 由四维向量

$$\boldsymbol{x}(0)=(ff',f(1-f'),(1-f)f',(1-f)(1-f')) \tag{3.20}$$

给出. 在以后的回合中, 这些概率根据两个参与者的反应规则变化. 我们用 p_{ij} 表示这一回合到下一回合状态由 i 变为 j $(i,j\in\{1,2,3,4\})$ 的概率. 因此, $\boldsymbol{x}(n)$ 根据转换规则

$$\boldsymbol{x}(n+1)=\boldsymbol{x}(n)\boldsymbol{P} \tag{3.21}$$

变为 $\boldsymbol{x}(n+1)$, 其中, $\boldsymbol{P}=(p_{ij})$ 是随机矩阵

$$\boldsymbol{P}=\begin{pmatrix} pp' & p(1-p') & (1-p)p' & (1-p)(1-p') \\ qp' & q(1-p') & (1-q)p' & (1-q)(1-p') \\ pq' & p(1-q') & (1-p)q' & (1-p)(1-q') \\ qq' & q(1-q') & (1-q)q' & (1-q)(1-q') \end{pmatrix}. \tag{3.22}$$

这就得到了一个马尔可夫链.

3.6 联　接

如果在状态 i 的概率 x_i 满足条件

$$x_1x_4=x_2x_3, \tag{3.23}$$

则两个参与者的行为是独立的. 事实上, x_1 是参与者 I 和参与者 II 都合作的概率. 参与者 I 合作的概率为 x_1+x_2, 参与者 II 合作的概率为 x_1+x_3. 独立性意味着

$x_1 = (x_1 + x_2)(x_1 + x_3)$, 当 $\boldsymbol{x} \in S_4$ 时, 等价于 $x_1x_4 = x_2x_3$. 在这种意义下, **联接** (linkage) $D = x_1x_4 - x_2x_3$ 衡量了两个参与者的相互依赖程度: $D = 0$ 意味着它们的行为是独立的.

直接计算可得

$$D(n+1) = \rho\rho' D(n), \tag{3.24}$$

如前, $\rho = p - q$ 和 $\rho' = p' - q'$. 事实上, 利用 (3.21), 只需要将 $D(n+1)$ 中的 $x_j(n+1)$ 替换为 $\sum_i x_i(n)p_{ij}$, 然后再比较乘积项 $x_i(n)x_j(n)$ 的系数. 由于 $p_{k1}p_{k4} = p_{k2}p_{k3}$ 和 $p_{1k}p_{4k} = p_{2k}p_{3k}$, 其中 $k = 1, 2, 3, 4$, 大部分系数可以抵消.

因此, **联接不平衡** (linkage disequilibrium) $D(n)$ 在初始回合为 0, 则保持为 0 (如果初始回合不为 0, 且至少其中一个反应规则是非确定的, 那么它会指数收敛到 0). 这证实了两个参与者每一回合的行为是独立的, 就像预期的那样.

3.7 合作程度

参与者使用反应策略就像彼此打乒乓: 如果参与者 II 以 y 的概率合作, 则参与者 I 在下一回合合作的概率为

$$\alpha(y) = py + q(1 - y) = q + \rho y. \tag{3.25}$$

因此, 如果把参与者 I 在 n 回合的**合作程度**记为 $c_n = x_1(n) + x_2(n)$, 则

$$c_{n+2} = q + \rho(q' + \rho' c_n) = A + uc_n, \tag{3.26}$$

其中, $u := \rho\rho'$, $A = q + \rho q'$ (即 $\alpha\alpha'(0)$). 等式 (3.26) 定义了一个从单位区间 (所有合作程度的集合) 到自身的仿射线性映射. 这个映射可以从初始回合开始迭代:

$$c_0 = f \mapsto c_2 = A + uf \mapsto c_4 = A + u(A + uf) = A(1+u) + u^2 f \mapsto \cdots. \tag{3.27}$$

由于

$$c_{2n} = A(1 + u + \cdots + u^{n-1}) + u^n f = \frac{A}{1-u} + u^n \left(f - \frac{A}{1-u}\right), \tag{3.28}$$

我们得到

$$c_{2n} = v + u^n(f - v), \tag{3.29}$$

其中

$$v := \frac{A}{1-u} = \frac{q + \rho q'}{1 - \rho\rho'} \tag{3.30}$$

正好是 $y \mapsto A + uy$ 的驻点. 一个类似的等式适用于 c_{2n+1} (用 $c_1 = q + \rho f'$ 替换 f). 因此, 合作程度 c_n 收敛到 v. 这同样适用于另一个参与者, 其合作程度收敛到 v'. 显然有 $\alpha(v') = v$ 等等. 只有当两个策略都有确定反应规则时合作程度才可能一直周期性震荡. 例如, TFT 参与者遇上一个使用 $(0,1,0)$ 的多疑的 TFT 参与者.

除了两个参与者相互的固定合作程度 v 与 v' 之外, 也可以考虑假想的合作程度 s 与 s', 这是两个采用相同策略的参与者的合作程度的极限. 一个 (f,p,q) 参与者对另一个 (f,p,q) 参与者达到的合作程度为

$$s := \frac{q}{1-\rho}. \tag{3.31}$$

有趣的是, $v - v'$ 与 $s - s'$ 同号 (且 $v - s'$ 与 $s - v'$ 符号也相同). 特别地, 如果合作程度的极限 v, v', s 和 s' 中有两个相等, 则四个都相等. 对此, 只需注意

$$v - s = \rho(v' - s). \tag{3.32}$$

这给出了一个把合作程度与反应规则联系起来的简单解释. 在 (p,q) 到 $(1,0)$ (TFT 规则) 连线上的每一个反应规则 (p',q'), 其与自身博弈时有相同的渐进合作程度, 因此与其他规则博弈时也相同. 如果反应规则 (p',q') 位于 (p,q) 到 $(1,0)$ 连线的上方, 则它有一个更高的渐进合作程度 (与自身, 或者与 (p,q)), 反之亦然.

3.8 收 益 值

我们将不考虑 $u^2 = 1$ 的情形 (只有当两个策略都有确定反应规则时才会出现).

由于两个参与者的决定是独立的, 使用策略 (f,p,q) 的参与者与使用策略 (f', p',q') 的参与者在第 n 回合的收益为

$$A(n) = Rc_nc'_n + Sc_n(1-c'_n) + T(1-c_n)c'_n + P(1-c_n)(1-c'_n). \tag{3.33}$$

特别地, 在捐赠博弈中上式简化为

$$A(n) = bc'_n - cc_n. \tag{3.34}$$

在无限重复的情形, 即 $w = 1$ 时, 这表示每一回合的平均收益为

$$(R-S-T+P)vv' + (S-P)v + (T-P)v' + P, \tag{3.35}$$

而对捐赠博弈为

$$bv' - cv. \tag{3.36}$$

这些表达式不依赖于初始的合作偏好 f 和 f'.

为了得到 $w<1$ 时捐赠博弈的总收益, 我们必须计算 $\sum w^n c_n$. 由等式 (3.29) 有

$$\sum w^n c_n = \sum w^{2n}[v+u^n(f-v)] + \sum w^{2n+1}[v+u^n(c_1-v)], \tag{3.37}$$

去掉因子 $[(1-w)(1-uw^2)]^{-1}$, 为

$$v(1-uw^2)-v(1-w)-vw(1-w)+f(1-w)+c_1w(1-w) \tag{3.38}$$

$$=vw^2(1-u)+(1-w)(f+qw+w\rho f')$$

$$=(q+\rho q')w^2+(1-w)(f+qw+w\rho f'). \tag{3.39}$$

合并有 f, f', q 和 q' 的项, 并令 $e:=(1-w)f+wq$, $e':=(1-w)f'+wq'$, 得到

$$\sum w^n c_n = \frac{e+w\rho e'}{(1-w)(1-uw^2)}. \tag{3.40}$$

因此, 每一回合的平均收益为

$$\frac{-c(e+w\rho e')+b(e'+w\rho' e)}{1-uw^2}. \tag{3.41}$$

3.9 失误的合作者、背叛者和回报者

假设捐赠意向以 ϵ 的概率失败, 拒绝意向以 $k\epsilon$ 的概率失败, 其中 $k\geqslant 0$. 在实施过程中将这两种失误区分开是有意义的, 特别是 $k=0$ 的情形. 例如, 一些原本要捐赠的参与者, 没有资金, 就不能实施. 但不太可能有这样的参与者, 不想付出却因太粗心而捐赠. 这样, 策略 AllC, AllD 和 TFT 分别由 $\boldsymbol{e}_1=(1-\epsilon,1-\epsilon,1-\epsilon)$, $\boldsymbol{e}_2=(k\epsilon,k\epsilon,k\epsilon)$ 和 $\boldsymbol{e}_3=(1-\epsilon,1-\epsilon,k\epsilon)$ 给出.

将表达式 (3.41) 应用到这三个策略中, 得到一个 3×3 的收益矩阵 $\boldsymbol{M}$, 初看起来 $\boldsymbol{M}$ 有些令人却步. 但它可以被极大地化简, 特别是当两个无条件策略的 ρ 值为 0 (即 $p=q$) 时. 我们再一次用到如下事实: 将 $\boldsymbol{M}$ 的每一列减去相应的对角线上的元素, 复制动力系统在 S_3 保持不变. 去掉乘积因子 $c(1-(k+1)\epsilon)$, 每一回合收益值的标准化矩阵为

$$\boldsymbol{M}=\begin{pmatrix} 0 & -1 & \delta\sigma \\ 1 & 0 & -\kappa\sigma \\ \delta & -\kappa & 0 \end{pmatrix}, \tag{3.42}$$

其中

$$\delta := w\epsilon, \quad \kappa := 1-w+wk\epsilon, \quad \sigma := \frac{b\theta-c}{c-c\theta}, \quad \theta = w(1-(k+1)\epsilon). \tag{3.43}$$

注意到, $\bar{P} = z(1+\sigma)P_z$. 利用

$$P_z - \bar{P} = P_z[1-(1+\sigma)z], \tag{3.44}$$

我们看到, 在 S_3 的内部, 当 $g := 1-(1+\sigma)z$ 为零时, $\dot{z}=0$ 成立. 容易知道, $g=0$ 对应着连接驻点 $F_{yz} := (0, 1-\hat{z}, \hat{z})$ 与 $F_{xz} := (1-\hat{z}, 0, \hat{z})$ 的轨线, 其中, $\hat{z} := (1+\sigma)^{-1}$. 在边界 $x=0$ 上, 背叛者和回报者处于双稳态竞争, 它们的吸引盆由 F_{yz} 分开. 在边界 $y=0$ 上, 回报者和 AllC 参与者在点 F_{xz} 稳定共存. 对于边界 $z=0$ 上的无条件参与者, 背叛者优于合作者.

在 S_3 内部我们得到常数运动

$$V := x^A y^B z^C[1-(1+\sigma)z], \tag{3.45}$$

其中, $A=\kappa/\theta$, $B=\delta/\theta$, $C=-1/\theta$ (注意: $A+B+C+1=0$). 内部驻点为

$$F = \frac{1}{1+\sigma(\kappa+\delta)}(\kappa\sigma, \delta\sigma, 1). \tag{3.46}$$

动力学行为如图 3.2 所示. $z=\hat{z}$ 是连接驻点 F_{xz} 与 F_{yz} 的水平轨线 (后者是纳什均衡). 在此线下方, 所有轨线收敛到 $y=1$, 背叛者胜利. 在水平轨线上方, 充满了围绕唯一驻点的周期轨线, 它们对应表达式 (3.45) 所给出的不变运动 V 的常数水平曲线. 时间平均对应驻点 F 的值. 这个驻点稳定, 但不渐进稳定. 我们注意到如果失误率 ϵ 充分小, 则背叛者在 F 的数量可以非常小. 另一方面, 如果 w 充分接近 1, AllD $(y=1)$ 的吸引盆可以任意小.

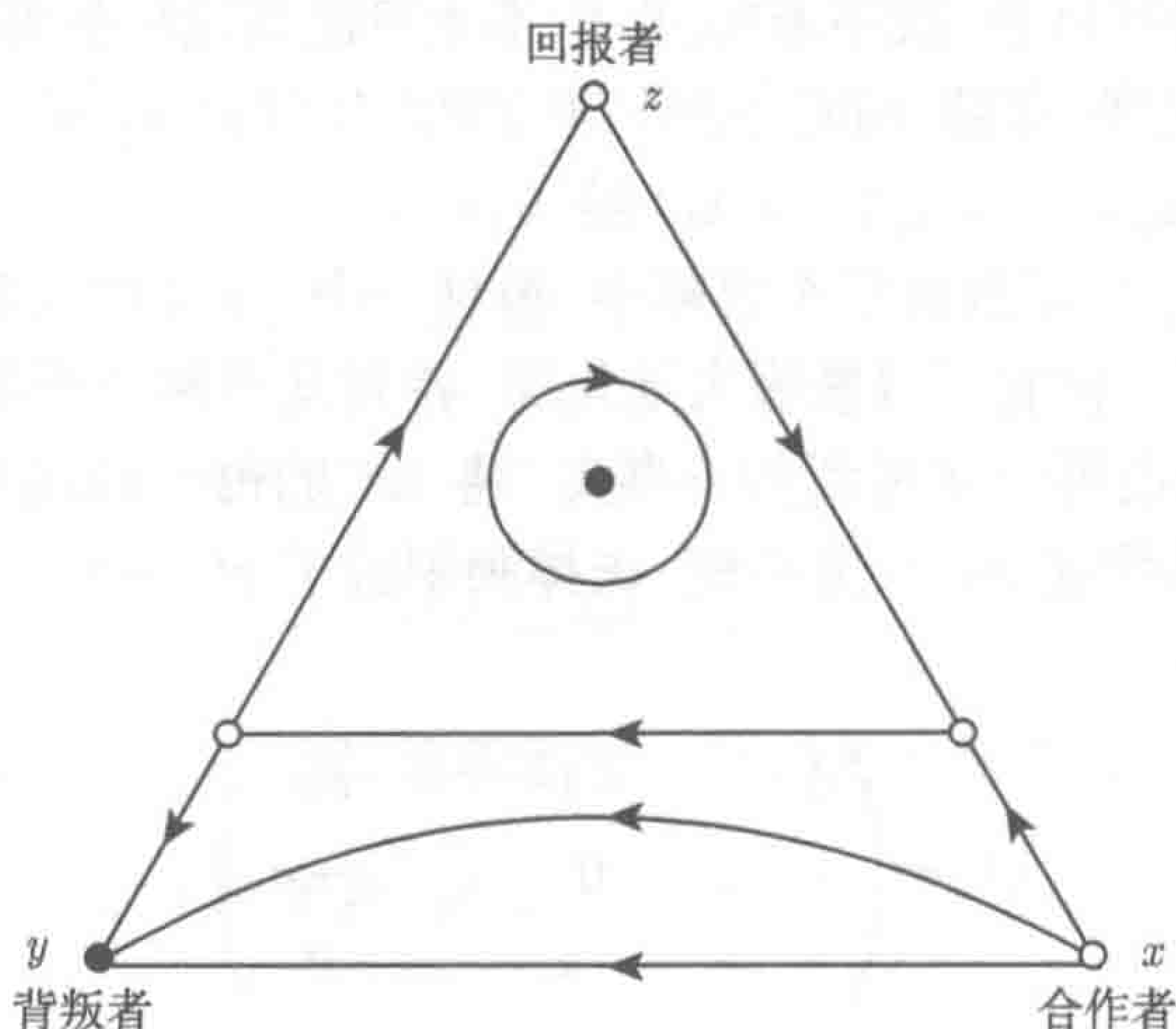


图 3.2　失误的合作者、背叛者和回报者. 如果 z 在临界值之下, 则背叛者胜利; 如果 z 在临界值之上, 三种策略共存, 它们的频度呈周期性震荡

3.10 极限情形

当 $w=1$ 时, 去掉乘积因子 $c(1-(k+1)\epsilon)$, 可得收益矩阵

$$\boldsymbol{M}=\begin{pmatrix}0 & -1 & \beta\\ 1 & 0 & -k\beta\\ \epsilon & -k\epsilon & 0\end{pmatrix}, \tag{3.47}$$

其中,

$$\beta:=\frac{1}{c}\left(\frac{b-c}{1+k}-\epsilon b\right). \tag{3.48}$$

如果 $k>0$ (即使本意是背叛, 也有捐赠的可能性), 则其动力学行为与图 3.2 相同, 对应分界线的 z 坐标为

$$\hat{z}:=\frac{c}{(b-c)}\frac{(k+1)\epsilon}{(1-(k+1)\epsilon)}. \tag{3.49}$$

若 $\epsilon\to 0$, 分界线与边界 $z=0$ 重合, 我们得到一个具有如下收益矩阵的系统:

$$\boldsymbol{M}=\begin{pmatrix}0 & -c & (b-c)/(1+k)\\ c & 0 & -k(b-c)/(1+k)\\ 0 & 0 & 0\end{pmatrix}. \tag{3.50}$$

这是一个石头—剪刀—布的博弈: AllD 被 TFT 取代, TFT 被 AllC 取代, AllC 又被 AllD 取代. S_3 内部唯一的驻点为 $F=\frac{1}{b}(k(b-c)/(k+1),(b-c)/(k+1),c)$. 复制动力学行为如图 3.3 所示.

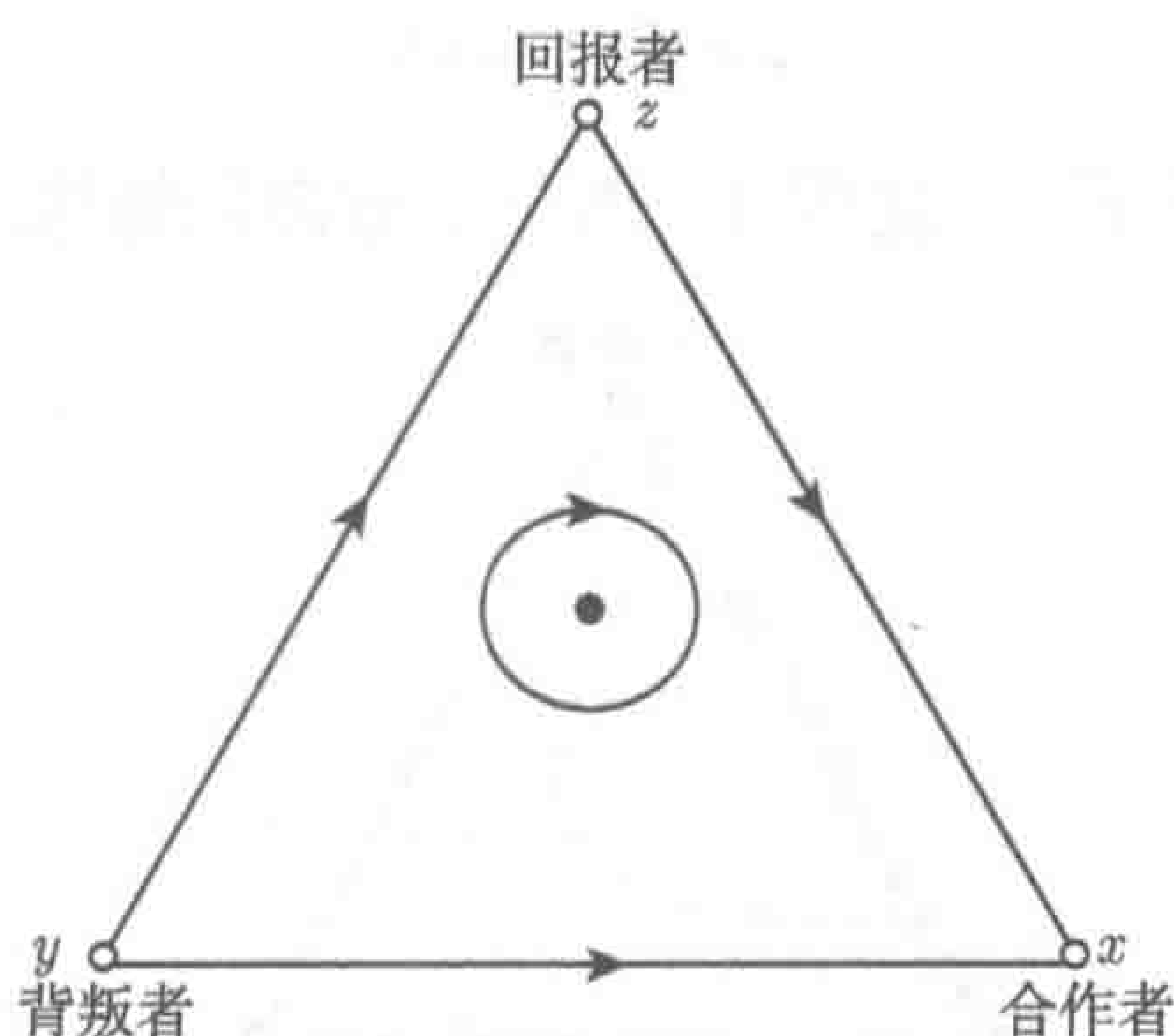


图 3.3 无限重复的捐赠博弈 ($w=1$), 如果计划的行为 (捐赠或是拒绝捐赠) 错误执行的概率大于零

另一方面, 如果首先考虑极限情形: $\epsilon=0$ ($w<1$), 就得到如图 3.1 所示的动力学行为. 如果又令 $w\to 1$, 得到图 3.4. 因此, $w\to 1$ 与 $\epsilon\to 0$ 不可交换.

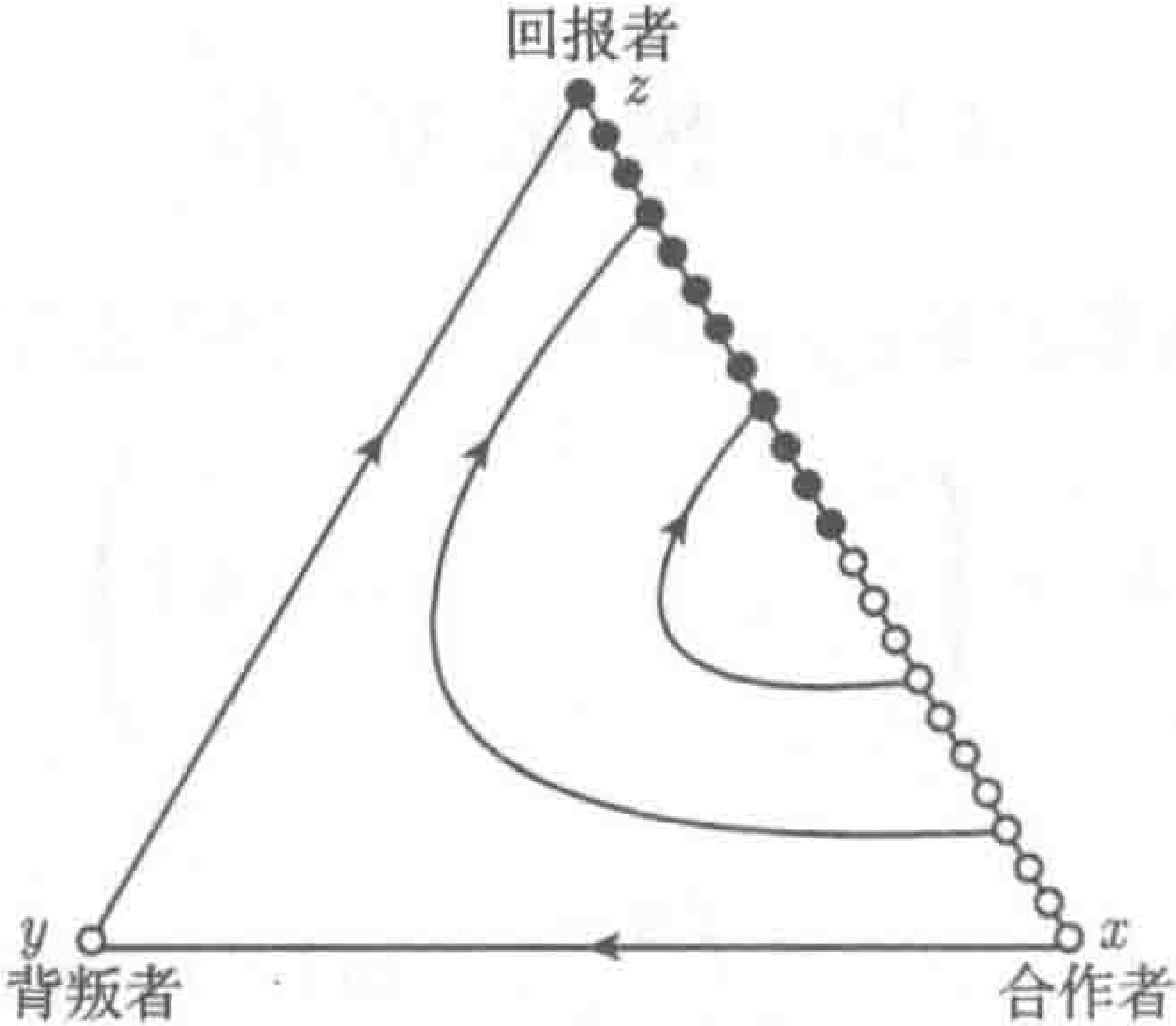


图 3.4　无失误的 ($\epsilon = 0$) 无限重复的 ($w = 1$) 捐赠博弈

现在令 $k = 0$, 即拒绝意向永远不会落空. 这是合理的. 在极限情形 $w = 1$ 时, 去掉收益矩阵的乘积因子 $c(1-\epsilon)$ 后得到

$$\boldsymbol{M} = \begin{pmatrix} 0 & -1 & \beta \\ 1 & 0 & 0 \\ \epsilon & 0 & 0 \end{pmatrix}, \tag{3.51}$$

其中, $\beta = [(1-\epsilon)b - c]/c$. 这将得到一个完全不同的相图. 边界 $x = 0$ 由驻点构成. 直观上看, 这是很显然的. 两个 TFT 之间的失误最终会导致双方在每一回合中相互背叛, 并且永远不可能通过失误再纠正回来. 因此, TFT 参与者每一回合的平均收益为 0. $z \leqslant \bar{z}$ 的驻点为纳什均衡, 其中

$$\bar{z} = c/[b(1-\epsilon)]. \tag{3.52}$$

其动力学行为如图 3.5 所示, 它是图 3.1 的一个有趣的镜像.

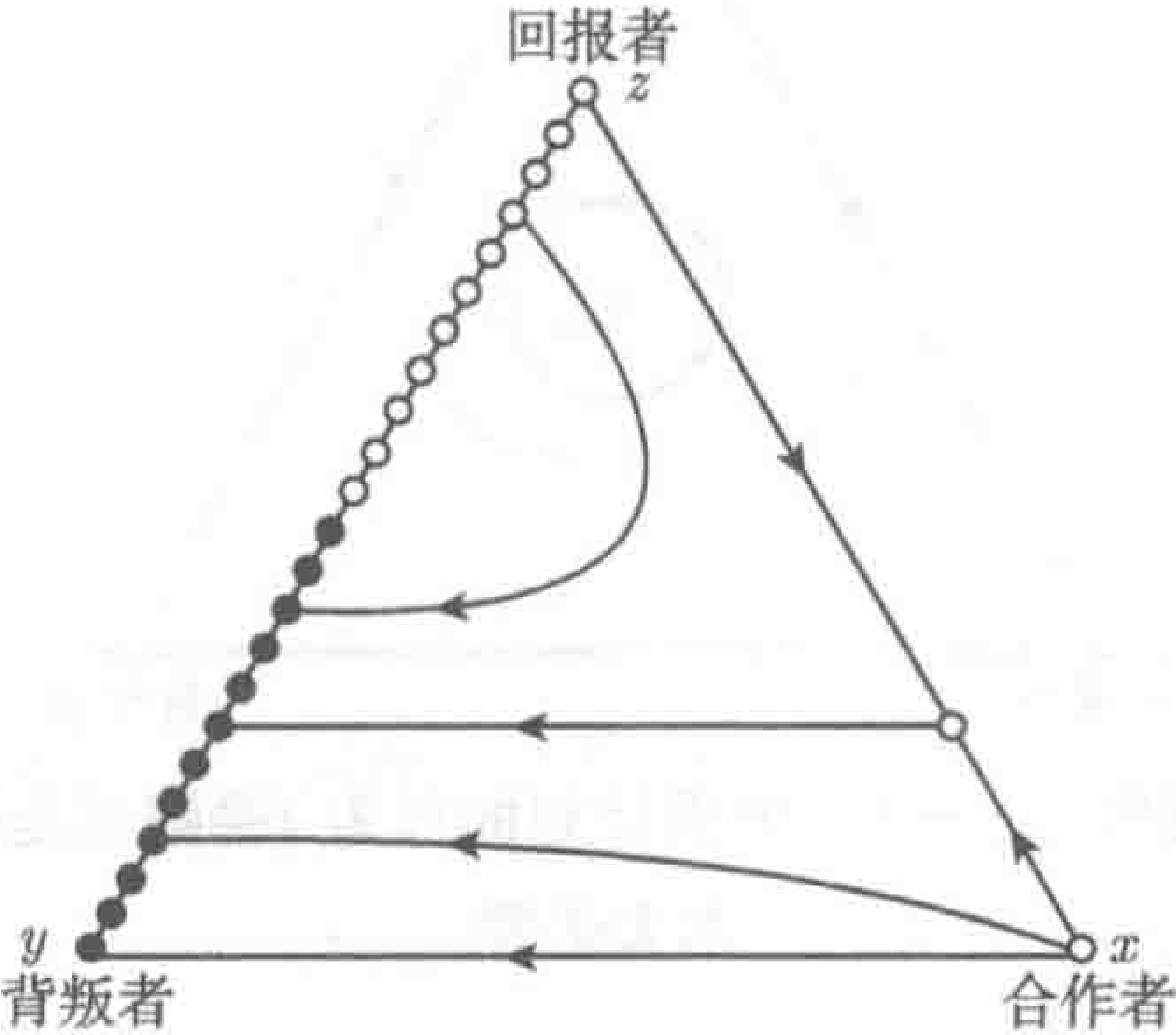


图 3.5　只有合作会被错误执行而背叛不会改变的无限重复的捐赠博弈的复制动力学行为. 从长远来看合作会消失

3.11 适应动力学

反应策略 (f,p,q) 形成一个连续统. 由三四种这类策略组成的异质群体可能出现极限环、异宿环或混沌震荡等复杂的动力学行为. 我们不考虑这一点, 只研究同质群体是怎样进化的.

更准确地, 我们假设常驻群体是同质的, 不时地, 有少量的其他类型进入. 这些外来者的收益要么好于常驻者, 要么差一些. 模仿将会发生, 这通常使得新类型或者固定下来, 或者被淘汰. 之后, 又有其他类型来试试它们的运气, 等等. 这样的极限情形 (极少创新, 极强模仿; 或者在生物学框架下表现为少突变, 强选择) 可以由一系列的同质群体来描述. 我们将描述*适应动力学* (adaptive dynamics), 它指向最有利的进化方向.

首先考虑 $w=1$ 的极限情形. 如果用 $\boldsymbol{n}:=(p,q)$ 表示常驻类型的反应规则, $\boldsymbol{n}':=(p',q')$ 表示稀少的侵入类型的反应规则, 我们必须判断侵入类型或是常驻类型哪个收益更好. 两种类型的参与者都必须与常驻类型接触 (由于侵入类型个体很少). 用 $A(\boldsymbol{n}',\boldsymbol{n})$ 表示使用策略 $\boldsymbol{n}'$ 的参与者对策略 $\boldsymbol{n}$ 参与者的平均收益. 因此, 当且仅当收益差 $A(\boldsymbol{n}',\boldsymbol{n})-A(\boldsymbol{n},\boldsymbol{n})$ 为正时类型 $\boldsymbol{n}'$ 能够侵入.

如 3.7 节, 用 s 表示一个 (p,q) 参与者对另一个 (p,q) 参与者的渐进合作程度, 分别用 v 与 v' 表示 (p,q) 参与者对 (p',q') 参与者对局时双方的渐进合作程度. 由等式 (3.34) 可得

$$A(\boldsymbol{n},\boldsymbol{n})=bs-cs \tag{3.53}$$

且

$$A(\boldsymbol{n}',\boldsymbol{n})=bv-cv'. \tag{3.54}$$

因此, 利用 $\rho=p-q$, 可得

$$A(\boldsymbol{n}',\boldsymbol{n})-A(\boldsymbol{n},\boldsymbol{n})=(v'-s)(b\rho-c). \tag{3.55}$$

直线 $\rho=c/b$, 即 $q=p-(c/b)$, 将反应规则 (p,q) 的正方形 $[0,1]^2$ 划为两个区域 (图 3.6), 即东南角 (包含 TFT 策略 $(1,0)$) 和其他部分. 如 3.7 节, $v'-s$ 符号的正负取决于 $\boldsymbol{n}'=(p',q')$ 是位于 $\boldsymbol{n}=(p,q)$ 到 $(1,0)$ 连线的上方还是下方. 因此, 如果 $\boldsymbol{n}$ 在东南角, 正好是那些更合作的策略 $\boldsymbol{n}'$ 能够侵入: 事实上, 如果 $s'>s$, 则 $v'>s$, 侵入者的收益大于常驻者的收益. 我们将这一区域记为*合作回报* (cooperation-rewarding) 区域. 相反地, 如果常驻类型采用一个不在合作回报区域内的策略 $\boldsymbol{n}=$

(p, q), 则合作度更低的策略都能侵入. 如果 $\boldsymbol{n}$ 在合作回报区域的边界上, 即满足 $\rho = c/b$, 则所有其他策略与 $\boldsymbol{n}$ 对局的收益和 $\boldsymbol{n}$ 与同类对局的收益相同.

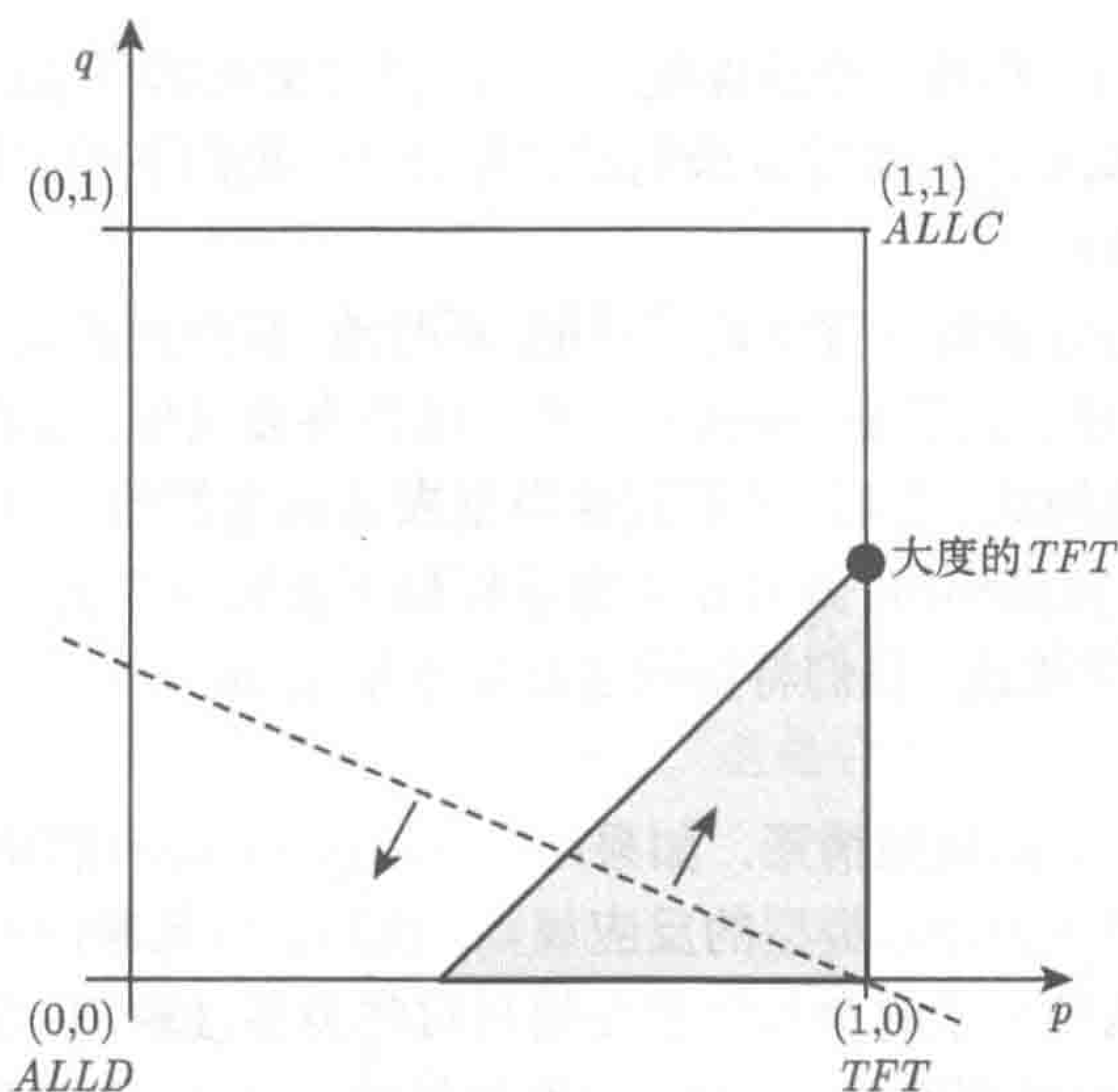


图 3.6　只合作回报区域 (灰色阴影部分) 是无限重复捐赠博弈反应规则 (p, q) 空间的一个子集. 箭头指向最有利的适应方向. 这个方向始终与连接规则 (p, q) 和 $(1, 0)$ 的直线正交

如果侵入者的策略 $\boldsymbol{n}'$ 接近常驻者的策略 $\boldsymbol{n}$, 我们可以用一阶泰勒展式

$$(p' - p)\frac{\partial A}{\partial p'}(\boldsymbol{n}', \boldsymbol{n}) + (q' - q)\frac{\partial A}{\partial p'}(\boldsymbol{n}', \boldsymbol{n}) \tag{3.56}$$

来逼近侵入者的收益差 $A(\boldsymbol{n}', \boldsymbol{n}) - A(\boldsymbol{n}, \boldsymbol{n})$, 其中函数 $\boldsymbol{n}' \mapsto A(\boldsymbol{n}', \boldsymbol{n})$ 的偏导数在 $\boldsymbol{n}' = \boldsymbol{n}$ 处取值. 于是, 我们将 (p, q) 的反应规则空间 $[0, 1]^2$ 中的适应动力学定义为

$$\dot{p} = \frac{\partial A}{\partial p'}(\boldsymbol{n}', \boldsymbol{n}), \quad \dot{q} = \frac{\partial A}{\partial q'}(\boldsymbol{n}', \boldsymbol{n}), \tag{3.57}$$

这里, 偏导数在 $\boldsymbol{n}' = \boldsymbol{n}$ 处取值. 对每一个同质状态 $\boldsymbol{n}$, 都将产生一个向量场, 它指向对侵入者最有利的方向. 简单的计算可以得到 $A(\boldsymbol{n}', \boldsymbol{n}) = bv - cv'$ 的导数. 这样, 有

$$\dot{p} \;=\; q\frac{b\rho - c}{(1-\rho)(1-\rho^2)}, \tag{3.58}$$

$$\dot{q} \;=\; (1-p)\frac{b\rho - c}{(1-\rho)(1-\rho^2)}. \tag{3.59}$$

因此, 在 $\boldsymbol{n} = (p, q)$ 的向量 $(\dot{p}, \dot{q})$ 与从 $\boldsymbol{n}$ 到 TFT 角 $(1, 0)$ 的向量正交. 当且仅当在合作回报区域, 向量朝上: 如果 p 增加 (感激), 则 q 也增加 (宽恕), 反之亦然.

这同样适用于一般的囚徒困境情形 ($w = 1$), 只是合作回报区域的形状不同: 在等式 (3.58) 与 (3.59) 中, $b\rho - c$ 这项由

$$(R - S - T + P)q\left(\frac{1+\rho}{1-\rho}\right) + (T - P)\rho + S - P \tag{3.60}$$

替代. 进化倾向没有指向 TFT: 这个策略是适应的枢轴, 而不是目标.

3.12 大度的 TFT

在合作回报区域的边界 $q = p - (c/b)$ 上的策略 $\boldsymbol{n}$ 有如下性质: 每个与 $\boldsymbol{n}$ 对局的策略 $\boldsymbol{n}'$ 有相同的收益 bq, 其最大值为 $b - c$, 在点 $(1, 1 - (c/b))$ 取得. 这个策略被称为**大度的** (generous) TFT (GTFT). 大度的 TFT 参与者在对手合作后总是合作, 但在对手背叛后并非总是背叛. 取而代之的是, 这样的参与者以确定的概率 $(b - c)/b$ 给予宽容.

GTFT 第一次出现在基于个体的计算机模拟, 如图 3.7 所示. 考虑一个大的虚构群体, 其参与者的策略是在 (p, q) 正方形中随机选取的. 这样, 初始群体不总是同质. 它可以由成百上千种不同类型的群体组成. 假设参与者随机相遇, 相互间进行大量的重复捐赠博弈. 进一步假设它们通过模仿更成功的参与者时而更新它们的策略. 很多策略将很快从群体中消失. 一般地, 初始策略集合中只有三个策略会起作用: 这些策略接近于 AllD, TFT 和 GTFT. 我们分别用 "AllD", "TFT" 及 "GTFT" 表示这些近似策略. 首先观察到的是朝向 "AllD" 的强烈趋势. 其他策略似乎无望地被远远超过. 但之后, "TFT" 不断增加, 并且取代 "AllD" (例如, 当 "TFT" 位于 "AllD" 和 TFT 的连线之下). 但这还不是最终的结局. 群体到达合作回报区域, 有更高 p 和 q 值的策略能再回来. 特别地, 更宽容的 "GTFT" 会取代严厉的 "TFT", 并在群体中固定下来. 惊人之处在于, 就其自身而言, "GTFT" 决不可能打败 "AllD", 需要 "TFT" 的催化. 这极像三个社会阶段的演变: 首先是 AllD 世界的 "狗咬狗", 然后是由 TFT 代表的 "复仇的规则", 最后是宽容但不是太宽容的 GTFT 阶段.

同样的结果对于 $w < 1$ 的捐赠博弈的适应动力学也成立. 此时, 初始回合的合作概率 f 是附加的特性. $\boldsymbol{n} = (f, p, q)$ 的适应动力学为

$$\dot{f} = \frac{bw\rho - c}{1 - w^2\rho^2}, \tag{3.61}$$

$$\dot{p} = \dot{f}\left(\frac{w}{1-w}\right)\left(\frac{e}{1-w\rho}\right), \tag{3.62}$$

$$\dot{q} = \dot{f}\left(\frac{w}{1-w}\right)\left(1 - \frac{e}{1-w\rho}\right), \tag{3.63}$$

其中, 如 3.8 节中的定义, $e=(1-w)f+wq$. 因为 $0<e<1-w\rho$, 所以这些分量同号. 因此, 又得到了一个合作回报区域. 直接的计算表明我们可以用示意的方式给出适应动力学:

$$\dot{f} = \frac{1-w}{1-w^2\rho^2}[A(\text{AllC},\boldsymbol{n}) - A(\text{AllD},\boldsymbol{n})], \tag{3.64}$$

$$\dot{p} = \frac{w}{1-w^2\rho^2}[A(\boldsymbol{n},\boldsymbol{n}) - A(\text{AllD},\boldsymbol{n})], \tag{3.65}$$

$$\dot{q} = \frac{w}{1-w^2\rho^2}[A(\text{AllC},\boldsymbol{n}) - A(\boldsymbol{n},\boldsymbol{n})]. \tag{3.66}$$

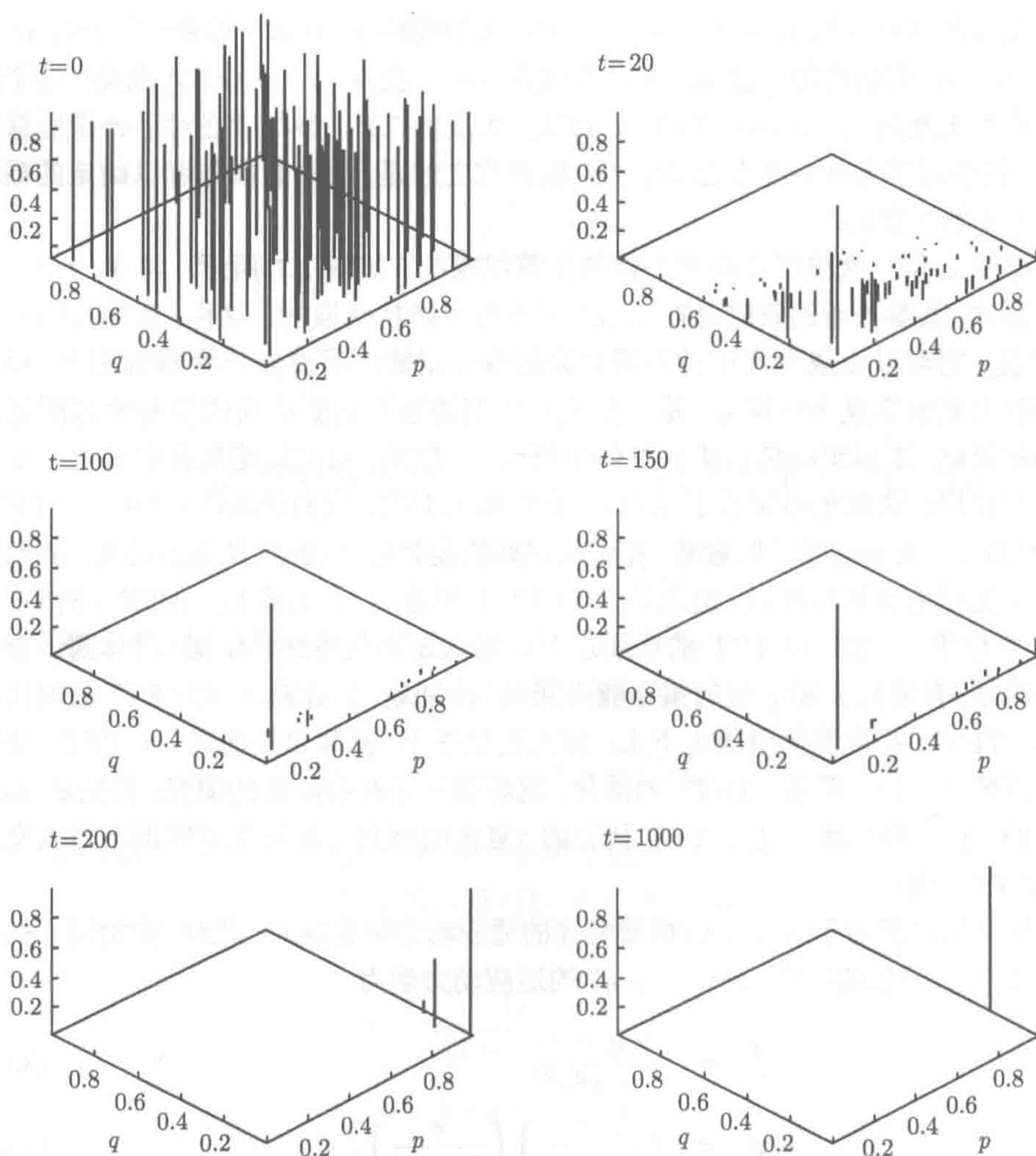


图 3.7　具有随机策略 (p,q) 的有限群体的进化. 首先, “AllD” 似乎胜利了, 然后是 “TFT”. 但最终 “GTFT” 获胜 (Nowak, Sigmund, 1992)

3.13 单步记忆策略

目前为止, 我们考虑了只依赖于对手前一回合行为的*反应策略* (reactive strategy). 但是, 一个合理的假设应该将参与者自己的行为也考虑进去. 如果对手的背叛是因为报复自己的背叛而不是利用自己的合作, 那么这种背叛更容易得到宽恕. 因此, 我们考虑随机策略 $(f, q_R, q_S, q_T, q_P) \in [0,1]^5$, 其中 f 如前, 是初始回合的合作倾向, q_R, q_S, $\cdots$ 分别表示前一回合收益为 R, S, $\cdots$ 时, 本回合的合作倾向.

假设使用策略 (f, q_R, q_S, q_T, q_P) 的参与者 I 与使用策略 $(f', q'_R, q'_S, q'_T, q'_P)$ 的参与者 II 对局. 这又一次涉及马尔可夫链; 每一回合的状态由参与者 I 获得的收益来表示. 转移概率由如下矩阵给出

$$\boldsymbol{Q} = \begin{pmatrix} q_R q'_R & q_R(1-q'_R) & (1-q_R)q'_R & (1-q_R)(1-q'_R) \\ q_S q'_T & q_S(1-q'_T) & (1-q_S)q'_T & (1-q_S)(1-q'_T) \\ q_T q'_S & q_T(1-q'_S) & (1-q_T)q'_S & (1-q_T)(1-q'_S) \\ q_P q'_P & q_P(1-q'_P) & (1-q_P)q'_P & (1-q_P)(1-q'_P) \end{pmatrix} \tag{3.67}$$

(一个参与者的 S 对应另一个参与者的 T).

四种状态的初始概率为向量:

$$\boldsymbol{x}(0) = (ff', f(1-f'), (1-f)f', (1-f)(1-f')), \tag{3.68}$$

我们用 $\boldsymbol{f}$ 来表示它. 在下一回合的概率为 $\boldsymbol{fQ}$, 在第 n 回合的概率为 $\boldsymbol{fQ}^n$. 对 $n \geqslant 1$, 概率不再位于联接平衡 (矩阵 $\boldsymbol{Q}$ 满足 $q_{k1}q_{k4} = q_{k2}q_{k3}$, 但不满足 $q_{1k}q_{4k} = q_{2k}q_{3k}$). 如果用 $\boldsymbol{g}$ 表示向量 (R, S, T, P), 则参与者 I 在第 n 回合的收益为

$$A(n) = \boldsymbol{g} \cdot \boldsymbol{fQ}^n. \tag{3.69}$$

如等式 (3.8) 所示, 当 $w < 1$ 时, 每一回合的平均收益为 $(1-w)\sum w^n A(n)$, 即

$$(1-w)\boldsymbol{g} \cdot \boldsymbol{f}(\boldsymbol{Id} - w\boldsymbol{Q})^{-1}, \tag{3.70}$$

其中 Id 是 4×4 单位矩阵. 对 $w = 1$, 我们必须另外计算. 如果矩阵 $\boldsymbol{Q}$ 是正规的, 即存在 m 使得 $\boldsymbol{Q}^m$ 的所有元素都是严格正的, 则存在特征值 1 及其唯一的左特征向量 $\boldsymbol{\pi} \in S_4$, 即 $\boldsymbol{\pi} = \boldsymbol{\pi Q}$. 元素 π_R, π_S, π_T 和 π_P 表示四种状态的平稳概率 (stationary probability), 则对每个初始状态 $\boldsymbol{f}$, 我们有

$$\boldsymbol{fQ}^n \to \boldsymbol{\pi}. \tag{3.71}$$

这时, 每一回合的平均收益为 $\boldsymbol{g} \cdot \boldsymbol{\pi}$, 在捐赠博弈中简化为

$$b(\pi_R + \pi_T) - c(\pi_R + \pi_S). \tag{3.72}$$

3.14　反应规则空间

当 $w=1$ 时, 可以忽略初始的合作概率, 只考虑反应规则 (q_R,q_S,q_T,q_P) 的空间. 这个单位方形由它的 16 个顶点张成, 即由四维向量 $(u_R,\cdots,u_P)$ 张成, 其中 u_i 取 0 或 1 决定于获得收益 $i\in\{R,S,T,P\}$ 之后策略选择合作还是背叛. 将这些策略记为 S_j, $u_Ru_Su_Tu_P$ 根据二进制给出, 其中 j 从 0 取至 15. 因此, AllD = $(0,0,0,0)$ 为 S_0, AllC = $(1,1,1,1)$ 为 S_{15}, TFT = $(1,0,1,0)$ 为 S_{10}, 等等. 如果计算 S_i 参与者与 S_j 参与者对局的转移矩阵 $\boldsymbol{P}$, 我们发现它一般是可约的, 当然就不是正规的. 矩阵中有很多元素为 0, 每一行只有一个元素不为 0, 因此, 平稳分布不是唯一确定的.

如果假设策略的实施会出现失误, 即策略以 ϵ 的概率被错误地执行, 则会出现不同的结果. 1 变为 $1-\epsilon$, 0 变为 ϵ. 策略 S_j 变为 $S_j(\epsilon)$. 例如 TFT, 即 $S_{10}=(1,0,1,0)$ 变为 $S_{10}(\epsilon)=(1-\epsilon,\epsilon,1-\epsilon,\epsilon)$, 等等.

容易计算 $S_i(\epsilon)$ 对 $S_j(\epsilon)$ 的收益. 相应的转移矩阵为 $\boldsymbol{Q}(\epsilon)$, 其元素是关于 ϵ 的二次多项式, 可以写为

$$\boldsymbol{Q}(\epsilon)=\boldsymbol{Q}+\epsilon\boldsymbol{Q}_1+\epsilon^2\boldsymbol{Q}_2, \tag{3.73}$$

其中, $\boldsymbol{Q}$ 为随机矩阵, 每一行刚好有一个 1, $\boldsymbol{Q}_1$ 和 $\boldsymbol{Q}_2$ 行之和为$\mathbf{0}$. 可以将 $\boldsymbol{Q}(\epsilon)$ 看成是矩阵 $\boldsymbol{Q}$ 的扰动, 将寻找 $\boldsymbol{Q}(\epsilon)$ 的左特征向量 $\boldsymbol{s}(\epsilon)$ 的问题视为扰动问题. 因此, 我们令

$$\boldsymbol{s}(\epsilon)=\boldsymbol{\pi}+\epsilon\boldsymbol{x}+\epsilon^2\boldsymbol{y}+\cdots, \tag{3.74}$$

其中随机向量 $\boldsymbol{\pi}$ 是未扰动特征值问题 $\boldsymbol{\pi}\boldsymbol{Q}=\boldsymbol{\pi}$ 的一个解, 而向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的分量之和为$\mathbf{0}$. 通过展开 $\boldsymbol{s}(\epsilon)\boldsymbol{Q}(\epsilon)=\boldsymbol{s}(\epsilon)$, 比较 ϵ 的幂次, 不仅得到 ($\epsilon\to 0$ 时) 收益的极限值 $\boldsymbol{\pi}$, 还可得到一阶项 $\boldsymbol{x}$.

例如, 考虑 $S_8=(1,0,0,0)$ 与 $S_{11}=(1,0,1,1)$ 的对局. S_8 也称为**冷酷** (grim), 因为它是 TFT 的冷酷变体, 除非双方上一回合相互合作, 否则它总选择背叛. 然而 S_{11}, 也称它为**坚定公平** (Firm But Fair, FBF), 是 TFT 宽容的兄弟, 如果上一回合两个参与者都背叛, 它选择合作.

这时

$$\boldsymbol{Q}=\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&0&0&1\\0&0&1&0\end{pmatrix} \tag{3.75}$$

表 3.1 具有行为误差的同步捐赠博弈

*	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
S_0	$\mathbf{0}$	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{3}$	b	$\frac{b}{2}$	b	0	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{2}$	b	$\frac{2b}{3}$	b
S_1	$-\frac{c}{2}$	$\mathbf{\frac{b-c}{2}}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	$\frac{b-2c}{5}$	$\frac{2b-c}{3}$	b	$\frac{3b-c}{4}$	$-\frac{c}{2}$	$\frac{2b-c}{3}$	$\frac{b-c}{3}$	$\frac{2b-c}{3}$	$\frac{2b-c}{4}$	b	b	b
S_2	0	$\frac{b-c}{3}$	$\mathbf{\frac{b-c}{4}}$	$\frac{b-c}{2}$	0	$\frac{2b-c}{3}$	0	$\frac{2b-c}{3}$	0	$\frac{b-c}{3}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	$\frac{2b-c}{4}$	$\frac{2b-c}{2}$	$\frac{4b-2c}{5}$	$\frac{2b-c}{2}$
S_3	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\mathbf{\frac{b-c}{2}}$	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$
S_4	$-\frac{c}{3}$	$\frac{2b-c}{5}$	0	$\frac{b}{2}$	$\mathbf{\frac{b-c}{4}}$	b	$\frac{b}{3}$	b	$-\frac{c}{3}$	$\frac{2b-c}{5}$	0	$\frac{b}{2}$	$\frac{3b-c}{6}$	b	$\frac{2b}{3}$	b
S_5	$-c$	$\frac{b-2c}{3}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$-c$	$\mathbf{\frac{b-c}{2}}$	$\frac{b-c}{2}$	$\frac{2b-c}{3}$	$-c$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-c}{3}$	$\frac{b-c}{2}$	b	b	b
S_6	$-\frac{c}{2}$	$-c$	0	$\frac{b-c}{2}$	$-\frac{c}{3}$	$\frac{b-c}{2}$	$\mathbf{0}$	$\frac{2b-c}{3}$	$-\frac{2c}{3}$	$-c$	$\frac{b-c}{2}$	$\frac{2b-2c}{3}$	$\frac{b-c}{2}$	$\frac{4b-3c}{5}$	$\frac{4b-2c}{5}$	$\frac{2b-c}{2}$
S_7	$-c$	$\frac{b-3c}{4}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$-c$	$\frac{b-2c}{3}$	$\frac{b-2c}{3}$	$\mathbf{\frac{b-c}{2}}$	$-c$	$-c$	$\frac{2b-2c}{3}$	$\frac{2b-2c}{3}$	$\frac{2b-3c}{4}$	$\frac{4b-3c}{5}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$
S_8	0	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{3}$	b	$\frac{2b}{3}$	b	$\mathbf{0}$	$\frac{3b-c}{5}$	0	$\frac{3b-c}{5}$	$\frac{3b-c}{6}$	$\frac{3b-c}{3}$	$\frac{3b-c}{4}$	$\frac{3b-c}{3}$
S_9	$-\frac{c}{2}$	$\frac{b-2c}{3}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	$\frac{b-2c}{5}$	$\frac{b-c}{2}$	b	b	$\frac{b-3c}{5}$	$\mathbf{b-c}$	$\frac{b-c}{2}$	$b-c$	$\frac{b-c}{2}$	$\frac{3b-2c}{3}$	$\frac{3b-c}{3}$	$\frac{2b-c}{2}$
S_{10}	0	$\frac{b-c}{3}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	0	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-2c}{3}$	0	$\frac{b-c}{2}$	$\mathbf{\frac{b-c}{2}}$	$\frac{2b-2c}{3}$	$\frac{b-c}{2}$	$b-c$	$b-c$	$b-c$
S_{11}	$-\frac{c}{2}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$-\frac{c}{2}$	$\frac{b-2c}{3}$	$\frac{2b-2c}{3}$	$\frac{2b-2c}{3}$	$\frac{b-3c}{5}$	$b-c$	$\frac{2b-2c}{3}$	$\mathbf{\frac{3b-3c}{4}}$	$\frac{2b-3c}{4}$	$b-c$	$b-c$	$b-c$
S_{12}	$-\frac{c}{2}$	$\frac{b-2c}{4}$	$\frac{b-2c}{4}$	$\frac{b-c}{2}$	$\frac{b-3c}{6}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{3b-2c}{4}$	$\frac{b-3c}{6}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{3b-2c}{4}$	$\mathbf{\frac{b-c}{2}}$	$\frac{5b-3c}{6}$	$\frac{5b-3c}{6}$	$\frac{2b-c}{2}$
S_{13}	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$-c$	$-c$	$\frac{3b-4c}{5}$	$\frac{3b-4c}{5}$	$\frac{b-3c}{3}$	$\frac{2b-3c}{3}$	$b-c$	$b-c$	$\frac{3b-5c}{6}$	$\mathbf{\frac{3b-3c}{4}}$	$\frac{3b-2c}{3}$	$\frac{3b-2c}{3}$
S_{14}	$-\frac{2c}{3}$	$-c$	$\frac{2b-4c}{5}$	$\frac{b-2c}{2}$	$-\frac{2c}{3}$	$-c$	$\frac{2b-4c}{5}$	$\frac{b-2c}{2}$	$\frac{b-3c}{4}$	$\frac{b-3c}{3}$	$b-c$	$b-c$	$\frac{3b-5c}{6}$	$\frac{2b-3c}{3}$	$\mathbf{b-c}$	$b-c$
S_{15}	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$\frac{b-3c}{3}$	$\frac{b-2c}{3}$	$b-c$	$b-c$	$\frac{b-2c}{2}$	$\frac{2b-3c}{3}$	$b-c$	$\mathbf{b-c}$

是可约的, 而

$$Q_1 = \begin{pmatrix} -2 & 1 & 1 & 0 \\ 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & -2 \\ 1 & 0 & -2 & 1 \end{pmatrix}. \tag{3.76}$$

由等式 $\boldsymbol{\pi Q} = \boldsymbol{\pi}$ 得$\pi_2 = 0, \pi_3 = \pi_4$, 即 $\pi = (1-2a, 0, a, a)$, 其中 a 是未知量. 由等式 $\boldsymbol{\pi Q}_1 + \boldsymbol{xQ} = \boldsymbol{x}$, 得 $a = \dfrac{2}{5}$, 因此 $\pi = \left(\dfrac{1}{5}, 0, \dfrac{2}{5}, \dfrac{2}{5}\right)$. 注意, 此时不需要 ϵ^2 项, 而有时是需要的.

表 3.1 列出了关于捐赠博弈的 16×16 矩阵 $\boldsymbol{A}$, a_{ij} 表示 $S_i(\epsilon)$ 参与者对 $S_j(\epsilon)$ 参与者的收益 (更准确地说是 $\epsilon \to 0$ 时的极限). 我们注意到一个明显的对称: 如果 $a_{ij} = xb - yc$, 则 $a_{ji} = yb - xc$.

如果常驻群体选择 S_0, 即 AllD, 则除了 $S_2 = (0,0,1,0)$, "冷酷" 策略 $S_8 = (1,0,0,0)$, 以及 TFT 策略 $S_{10} = (1,0,1,0)$ 以外, 其他策略都不能侵入. 在没有其他策略的情形下 S_2 被 S_{10} 占优. TFT 能战胜 AllD, 但 TFT 能被更宽容的策略取代, 如 S_{15}, 即 AllC; 而 AllC 又会被 AllD 取代. 但这个循环趋势能被 S_9 打破. 这个策略优于 S_{10}, S_2 以及 S_8 (如果 $b > 3c$). 只要 $b > 2c$, 即成本 — 收益比小于 1/2, 它就不能被 AllD 侵入.

注意到 S_9 是唯一不能被任何其他 S_i 策略侵入的策略 (当 $b > 3c$ 时). 并且 S_9 与自己博弈也很好: 一个 S_9 群体的参与者获得 $b - c$ 的收益, 是同质群体中博弈收益最好的. 只有 S_{14} 与 S_{15} 的收益与它一样好, 但它们容易被 S_1 或 S_0 取代.

3.15　胜—保持, 败—改变

$S_9 = (1,0,0,1)$ 这个策略比较令人费解, 我们姑且称它为巴甫洛夫 (Pavlov) 策略。如果两个参与者都使用 Pavlov 策略相互博弈, 通常都会合作. 例如, 参与者 II 由于失误选择了背叛, 那么在下一回合两个参与者都会背叛, 从那以后又重新开始相互合作, 像一对老夫妇争吵后一样 (图 3.8). 更进一步, 如果 Pavlov 参与者与 AllC 对局, 它会心安理得地利用对手. 在第一次偶然地背叛之后, 它会一直背叛直到再一次出错. 这是保护群体最后不被背叛者侵入的重要性质. 例如, 一个 TFT 群体将很快被 AllC 参与者取代, 而 AllC 参与者非常容易被 AllD 剥削.

当且仅当前一回合与对手做法相同, 策略 $\boldsymbol{S_9}$ 选择合作. 这个机制暗示了一个性质, 如图 3.9 所示. 如果获得正的收益 (奖励 $b - c$, 或者是更好的诱惑 b), 这个策略会有效地重复前一回合的做法. 如果收益是非正的 (两个参与者都背叛, 收益为 0, 或是傻瓜的收益 $-c$), 它会转变到其他策略. 这是最简单和容易想到的学习机制,

动物训练员和家长都熟知这一点. **胜—保持, 败—改变** (Win-Stay, Lost-Shift) 是广泛传播的动物行为准则.

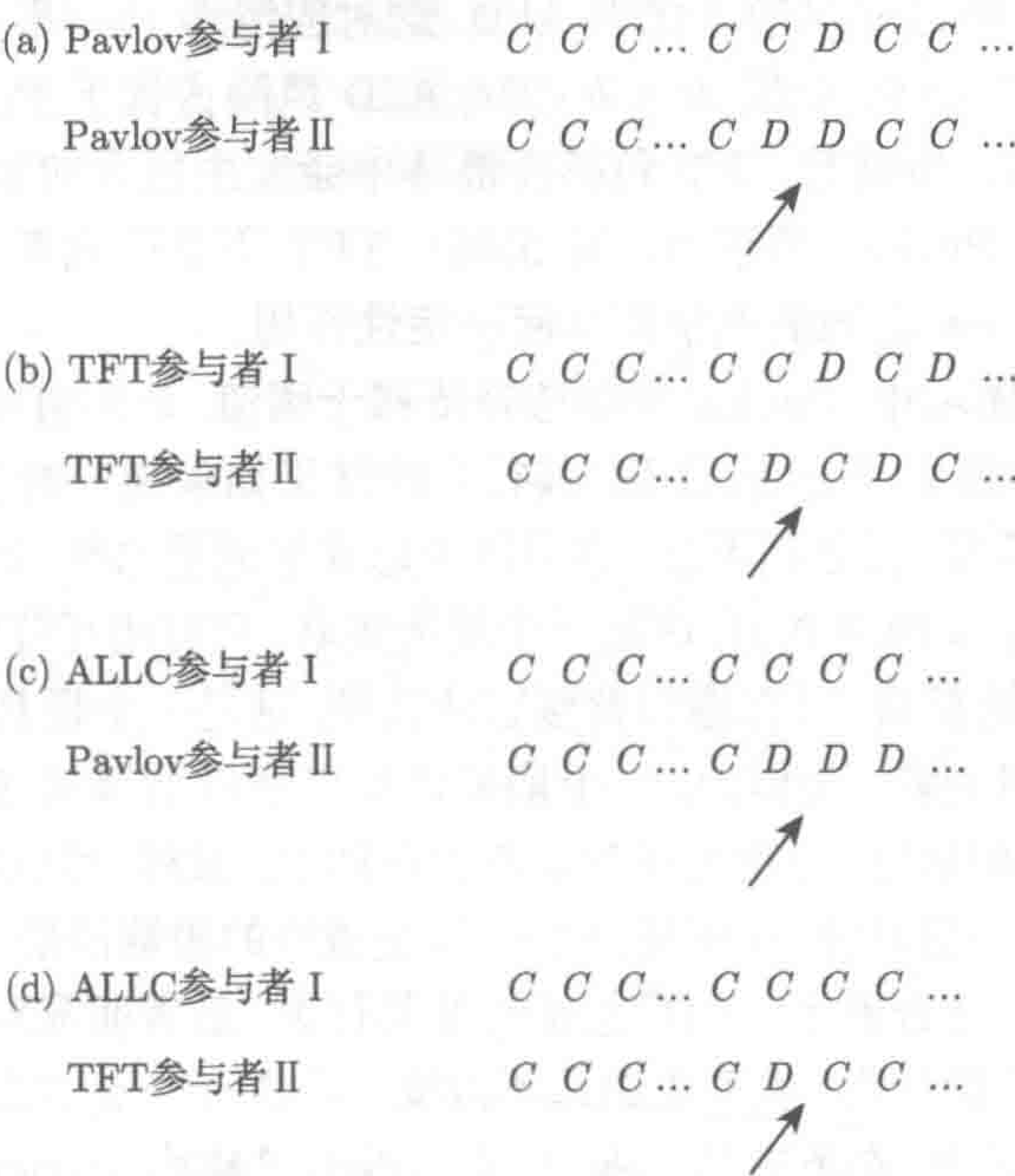


图 3.8 重复囚徒困境中失误的影响. 箭头表示每次运行错误执行的行为

Pavlov 参与者 I	C	C	D	D
参与者 II	C	D	C	D
	↓	↓	↓	↓
I 的收益	R	S	T	P
I 的下一回合策略	C	D	D	C

图 3.9 Pavlov 是一个胜—保持, 败—改变的策略. 在获得更大的收益值 R 和 T 之后, Pavlov 重复前面的成功做法. 在获得小的收益值 P 和 S 后, Pavlov 参与者改变为其他的做法

条件 $b > 2c$ 表明 Pavlov 不会被 AllD 占优, 但这两种策略处于双稳的争中. 条件 $b > 3c$ 表明 Pavlov 风险占优.

在这种背景下考虑有限群体是有趣的. 我们考虑两种情形: (a) $b = 5c/2$ 以及 (b) $b = 4c$, 群体规模 $M = 100$, 选择强度 $s = 1/10$. 假设在不可转移的情形 (很少的创新率 μ, 见 2.17 节). 在 S_0 = AllD 和 S_{10} = TFT 组成的群体中 TFT 占优. (由表达式 (2.90) 给出的平稳分布, TFT 在情形 (a) 有 97%, 在情形 (b) 有 99%.)

这表明 TFT 优于 AllD. 但如果群体中出现 AllC $= S_{15}$, 则平稳分布被 AllD 占优 (情形 (a)64%, 情形 (b) 66%). 现在考虑由策略 $S_9 =$ Pavlov, AllC 和 AllD 构成的群体. 如果 $b > 3c$, Pavlov 风险占优于 AllD, 这对应情形 (b), 事实上我们看到, 平稳分布中 Pavlov 占 90%. 如果 $2c < b < 3c$, AllD 风险占优于 Pavlov, 在平稳分布中有 80%的背叛者. 如果把 TFT 包括在群体中会发生巨大的变化: 这种情形下, 情形 (a) 有 50%的 Pavlov (情形 (b) 有 95%). TFT 不是胜利者, 但可以充当造王者—对 AllD 与 Pavlov 之间的竞争结果起决定性作用.

在一般的囚徒困境中, Pavlov 策略选择依赖于阈值, 该阈值将两个好的结果 T 和 R 与两个差的结果 P 和 S 区分开. 其行为的结果如果高于**期望水平** (aspiration level), 则行为会被重复, 反之则不会. 也可以考虑其他期望水平. 比如更具野心的只满足于 T 的策略 $S_1 = (0,0,0,1)$, 这是一个霸道策略, 只有相互背叛后才合作. 一旦遇到傻瓜它都无情地背叛, 只有遇到背叛它才改变. 这是一个极具野心的胜—保持, 败—改变策略, 它将失败. 类似地, 一个期望水平更适度的策略 $S_8 = (1,0,0,0)$ (P 和 S 之间), 它做得相当好, 特别是在背叛者的群体中. 最后, 可以将 $S_3 = (0,0,1,1)$ 和 $S_{12} = (1,1,0,0)$ 看做是胜 — 保持, 败 — 改变策略的极端情形. 前者从来不满足于任何结果, 从这一回合到下一回合总是改变其行为. 后者明显对任何结果都感到满意, 总是重复自己的行为, 除非出错从不改变. 显然, 有 "适当的" 期望水平是很重要的. 在这个意义上, 在所有胜—保持, 败—改变策略中 Pavlov 是最平衡的, 它在重复的雪堆博弈中也做得很好, 这由矩阵 (3.4) 所描述.

3.16　自　动　机

有确定反应规则的单步记忆策略容易由有限自动机实现. 例如, Pavlov 可以由仅有两种内部状态的自动机实现, 这两种内部状态用**相同** (Same) 和**相异** (Diff) 表示. 如果在前一回合, 参与者收益为 R 或 P (即两个参与者同时合作或者同时背叛), 则自动机的状态为 Same, 否则状态为 Diff. 在每种状态下, 自动机指定下一回合的行为, 即合作或是背叛. 因此, 对 Pavlov 的分析可以容易地通过有向图来完成, 见图 3.10. 有向图的节点为参与者的状态 Same 和 Diff. 每个节点有两个有向箭头离开, 一个实线, 一个虚线. 实线箭头描述参与者采用由 Pavlov 策略指定的行动 (即在 Same 时合作, 在 Diff 时背叛) 时状态的转变. 虚线箭头描述参与者采用其他行为时状态的转变. 在这两种情形下, 都假设对手使用 Pavlov 策略. 沿着每个箭头, 我们可以看到参与者相应的收益. 显然, 如果 $2R > T + P$ (当 $w < 1$ 时, $R + wR > T + wP$), 最好遵循实线箭头. 对于捐赠博弈, 可以简化为熟悉的条件 $b > 2c$ (对应的 $w(b-c) > c$). 如果这个条件满足, 则在与 Pavlov 参与者对局时, 选择 Pavlov 规则总是最好的. 在一个 Pavlov 参与者的群体中, 最好与它们保持

一致. 所有的单步记忆策略也可以用类似的图来研究. 一般地, 这要比 Pavlov 更复杂, Pavlov 中的两个参与者总是处于相同状态. 但四个状态 (C, C), (C, D), (D, C), (D, D) 足以描述自动机 (第一项描述参与者的行为, 第二项描述对手的行为). 图 3.11 描述了对手使用策略 TFT 时所产生的各种现象. 如果在节点处的参与者也采用 TFT 策略指定的行为, 则用离开节点的实线箭头描述状态的转变, 否则用虚线箭头描述.

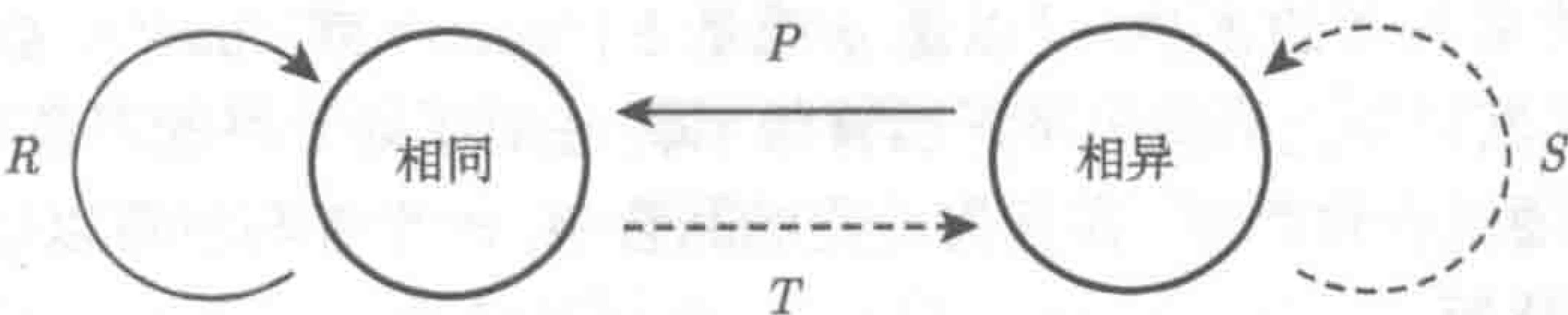


图 3.10 由两状态自动机描述的 Pavlov 策略. 假设对手使用 Pavlov 策略. 实线箭头和虚线箭头分别描述了参与者遵循 Pavlov 和偏离 Pavlov 策略时状态的转变. 若 $2R > T + P$ (捐赠博弈中为 $b > 2c$), 则 Pavlov 策略是自身的最佳响应

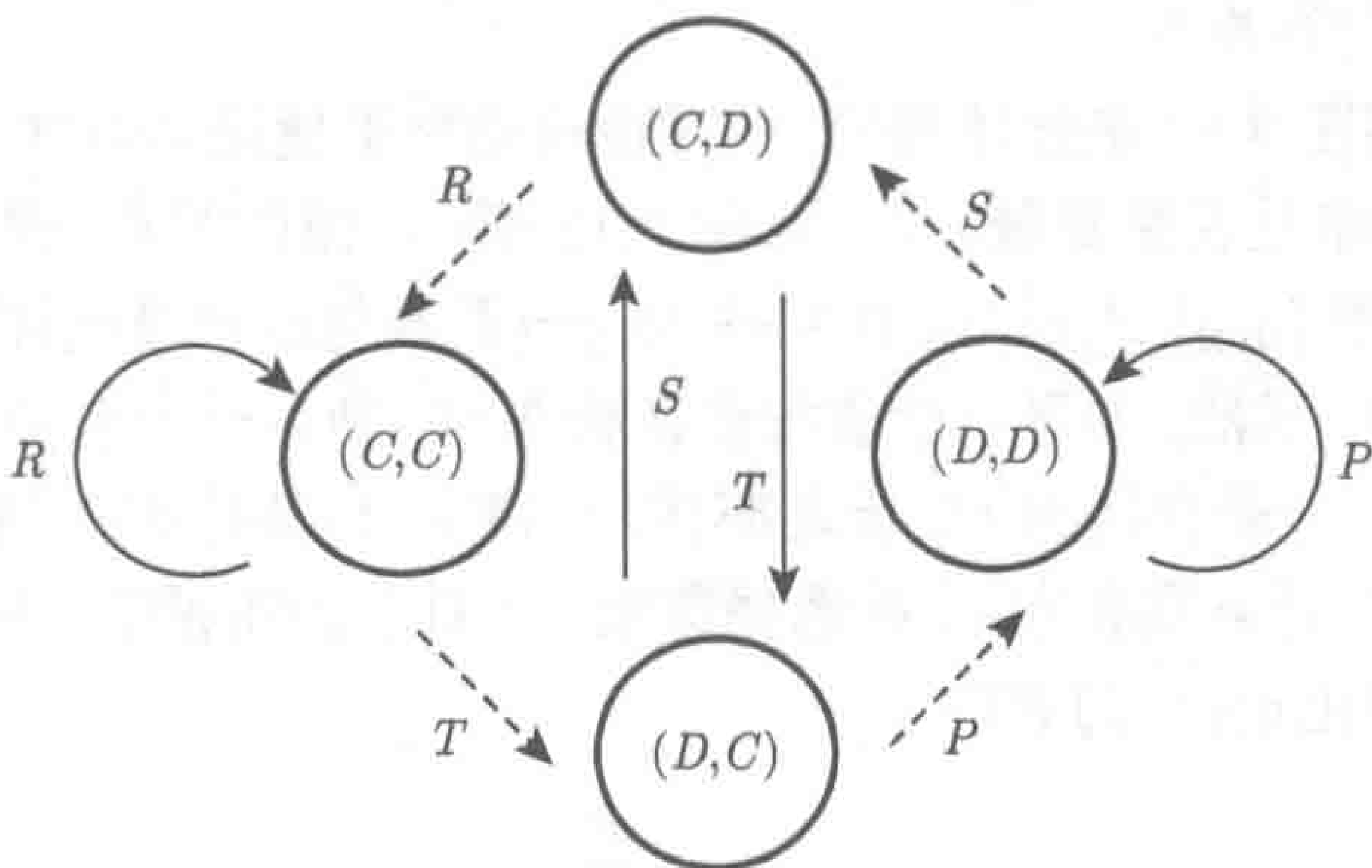


图 3.11 由有限自动机描述的 TFT 策略. 当 $2R > T + S$ 时, TFT 不是自身的最佳响应, 而捐赠博弈总是满足条件 $2R > T + S$

注意到在捐赠博弈中, TFT 不是自身的最佳响应. 在状态 (C, D), 与一个 TFT 参与者对局最好的做法是合作 (到达 (C, C) 状态). 然而 TFT 要求参与者背叛, 这使两个 TFT 参与者被锁定在单方面背叛的死循环之中. 这样, 每一回合的收益为 $(b-c)/2$, 这比到达节点 (C, C) 的收益 $b-c$ 要小. 当然, 两个 TFT 参与者将从节点 (C, C) 开始, 且在此状态下, TFT 给出了正确的选择. 但是, 如果出现失误会将状态引向节点 (C, D), 这显出了 TFT 致命的弱点. 更一般地, 如果 $2R > T + S$, 则这个策略不是自身的最佳响应.

这种方法可以推广到记忆依赖于前两回合或者是前 N 回合的所有策略. 但马上会看到, 一些非常简单的由有限自动机实现的策略不能被描述为指定回合数的策略.

3.17　悔 悟 的 TFT

一个有趣的例子是 CTFT (悔悟的 TFT). 设想一个未正确实施合作行为的 TFT 参与者, 意识到错误后, 会顺从地接受对手下一回合报复性背叛. 此时, 毫无意义的单边背叛的交替报复可以避免, 而相互合作可以重新开始. 为了建立这种模型, 我们引进参与者的*声誉*, 可以是 g 或是 b ("good" 或 "bad"). 参与者以好的声誉开始, 保持其声誉直到做出不正当背叛 (即, 在对手处于好的声誉时背叛对手). 好的声誉可以通过合作恢复. 在任何给定的回合中, 一个参与者可以合作、正当背叛或者不正当背叛.

CTFT 是这样一种策略, 只有在自己有好的声誉而对手没有时选择背叛. 这意味着参与者只有在被挑衅时才背叛, 否则不会背叛. 因此, 只要不失误, 重复囚徒困境博弈的两个 CTFT 参与者会一直合作. 失误之后它们会重新开始合作, 接受对手报复性背叛而不觉得吃亏.

如前, 我们用图 3.12 来描述博弈, 同时检验当对手使用 CTFT 策略时, 参与者也使用 CTFT 策略是否总是最好的, 如图 3.12 所示. 图的节点, 即博弈的状态, 为 A, B 和 C. A 对应 (g,g) 或 (b,b), B 对应 (b,g) (即参与者声誉是坏的, 而对手是好的), C 对应 (g,b). 显然, 如果一个参与者在状态 A, 则另一个也在, 如果一个参与者在状态 B, 另一个则在状态 C, 反之亦然. CTFT 的规则要求, 参与者在状态 A 或 B 时选择合作, 只有在状态 C 时选择背叛. 由对应的图说明: 与一个 CTFT 参与者对局时, 最好也选择 CTFT.

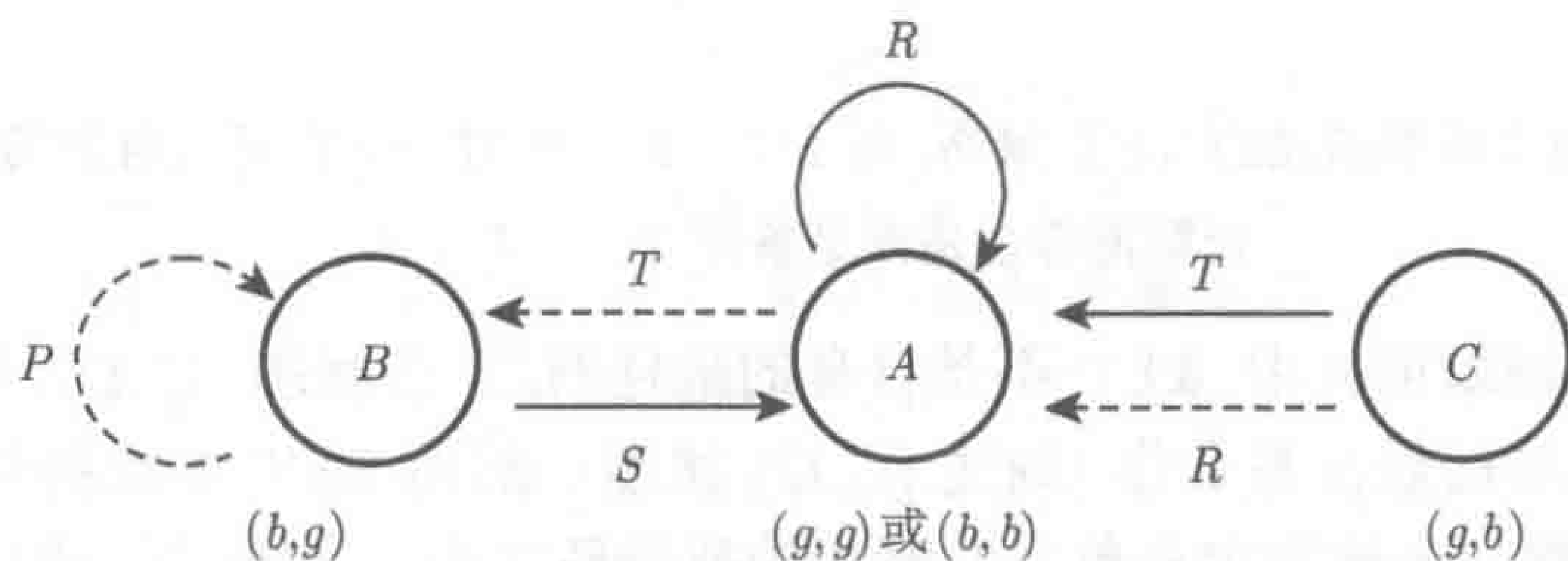


图 3.12　由三状态自动机描述的 CTFT 策略总是自身的最佳响应

CTFT 策略不能被描述为单步记忆策略. 它也不是一个只依赖于给定的前 N 回合结果的规则. 事实上, 假设在前 N 回合我们观察到一系列相互背叛. 如果我们不知道在这 N 回合之前发生了什么, 我们无法判断两个参与者中哪一个有好声誉, 哪一个有坏声誉. 因此, 我们不能确定 CTFT 参与者下一回合应该做什么.

在捐赠博弈中将 CTFT 与 Pavlov 进行比较是有趣的. CTFT 总是自身的最佳响应, Pavlov 只有当 $b > 2c$ 时才是. 两个策略遇到同类都选择合作, 偶然失误之后

能够很容易地恢复合作. CTFT 有一个巨大的优势, 即在侵入 AllD 参与者群体时与 *TFT* 同样有效; 正如我们看到的, 在这一点上 Pavlov 无法做到, 它需要一个报复性的策略为之铺路. 另一方面, 在一个由 CTFT 主导的社会中, 无条件的利他主义者收益与 CTFT 一样好, 因此可以通过中性漂变传播, 最后使得 AllD 侵入并破坏它们的合作机制. 相反的, 一个 Pavlov 参与者的社会不会允许 AllC 参与者传播. 一旦出现失误, AllC 参与者将会被彻底剥削.

当 Pavlov 表现不好, 即 $c < b < 2c$ 时, 另一个基于声誉的被称为悔恨者 (Remorse) 的策略, 其表现与 CTFT 一样好, 悔恨者策略只有在自己具有坏声誉或两个参与者前一回合都合作时才合作. 单边失误之后, 两个悔恨者分别背叛两次. 如果一个悔恨者与一个 Pavlov 参与者对局, 两者在每一回合的平均收益都为 $5(b-c)/7$.

3.18 感知的误差

CTFT 也有其阿基里斯之踵 (Achilles heel). 目前为止, 我们只考虑了实施过程中的失误. 在感知上出错会怎么样呢? 此时, 参与者自认为有好的声誉, 而对手认为其有坏的声誉. 在这种情形下, 两个 CTFT 参与者会无情地相互背叛, 并且都认为自己的背叛是正当的, 而对手的不是. 另一方面, 如果两个 Pavlov 参与者之间出现感知上的错误, 在通常相互惩罚的回合之后合作会顺利地恢复.

对于单步记忆策略, 如果行为实施失误的概率为 ϵ, 则在某回合收益为 R 之后合作倾向 q_R 由 $(1-\epsilon)q_R + \epsilon(1-q_R)$ 替代, 等等. 因此 "修正项"

$$\epsilon(1-2q_R, 1-2q_S, 1-2q_T, 1-2q_P) \tag{3.77}$$

必须被加到反应规则 (q_R, q_S, q_T, q_P) 中. 如果失误影响了对手的感知 (即参与者混淆了 R 与 S, T 与 P), 则 q_R 变为 $(1-\nu)q_R + \nu q_S$, 等. 而修正项为

$$\nu(q_S - q_R, q_R - q_S, q_P - q_T, q_T - q_P). \tag{3.78}$$

如果以 μ 的概率失误影响参与者自己的感知 (即参与者混淆了 R 与 T, S 与 P), 则修正项为

$$\mu(q_T - q_R, q_P - q_S, q_R - q_T, q_S - q_P). \tag{3.79}$$

如果两种类型的感知都出现失误, 则 TFT 的反应规则 $(1,0,1,0)$ 变为 $(1-\nu, \nu, 1-\nu, \nu)$, Pavlov $(1,0,0,1)$ 变为 $(1-(\nu+\mu), \nu+\mu, \nu+\mu, 1-(\nu+\mu))$, 而无条件策略 AllD 与 AllC 不受影响. 当 $w=1$, 而极限 $\epsilon \to 0$ 时, 一个使用冷酷策略 $(1,0,0,0)$ 的 $\boldsymbol{S_8}$ 参与者与使用 $(0,0,1,0)$ 的 S_2 参与者对局时, 执行失误产生 $(2P+2S+T)/5$ 的收益, 感知失误产生 $(S+T)/2$ 的收益.

因此, 考虑不同类型的失误是重要的. 例如, 我们可以作一个合理的假设 (3.10 节): 只有合作时才可能出错, 而背叛时不会. 此时, $S_8(\epsilon)$ 参与者与 $S_{11}(\epsilon)$ 参与者对局时的平稳分布为 $(0,0,1/2,1/2)$, 而不是 $(1/5,0,2/5,2/5)$ (3.14 节), 因此收益为 $b/2$ 而不是 $(3b-c)/5$. 容忍的坚定公平 (FBF) 参与者每次的第二回合试图恢复合作都是徒劳的. 这时 TFT 参与者对自身的收益就更差了 (为 0), 但 Pavlov 群体的收益保持不变. 再一次, 当 $b>2c$ 时, Pavlov 群体不能被侵入.

即使在只有三四个内部状态的自动机中, 也存在使人迷惑的策略种类. 似乎很难确定哪一个策略将会被进化选中. 基于模拟的个体展示出很多偶然性, 而很少提供稳健性预测. 我们遭遇一个复杂性壁垒. 另一方面, 用我们情绪来解释自动机的 "内部状态" 看起来是诱人的, 这些情绪包括被激发的愤怒以及偏离规范的愧疚等等.

3.19 触发器和均衡器

关于**重复博弈的大众定理** (folk theorem on repeated games) 是一批结果. 最简单的情形, 对处于无限重复捐赠博弈中的两个参与者 I 和 II, 这个定理表明, 任何满足 $0 \leqslant P_{\rm I}, P_{\rm II} \leqslant b-c$ 的收益对 $(P_{\rm I}, P_{\rm II})$ (每个回合), 可以由策略的一个纳什均衡对得到. 两个参与者只需要简单地遵循**触发器策略** (trigger strategies): 这意味着采取一系列明确的指向 $(P_{\rm I}, P_{\rm II})$ 的行为, 一旦对手偏离, 便转向无情的、无限的背叛. 显然, 对手没有偏离动机: 没有比遵循确定的行为更好的选择. 事实上, 任何使得 $P_{\rm I}$, $P_{\rm II}$ 为正, 且 $(P_{\rm I}, P_{\rm II})$ 在由 $(0,0)$, $(b,-c)$, $(-c,b)$ 和 $(b-c,b-c)$ 张成的凸包上的收益值对都可以得到, 如图 3.13 所示.

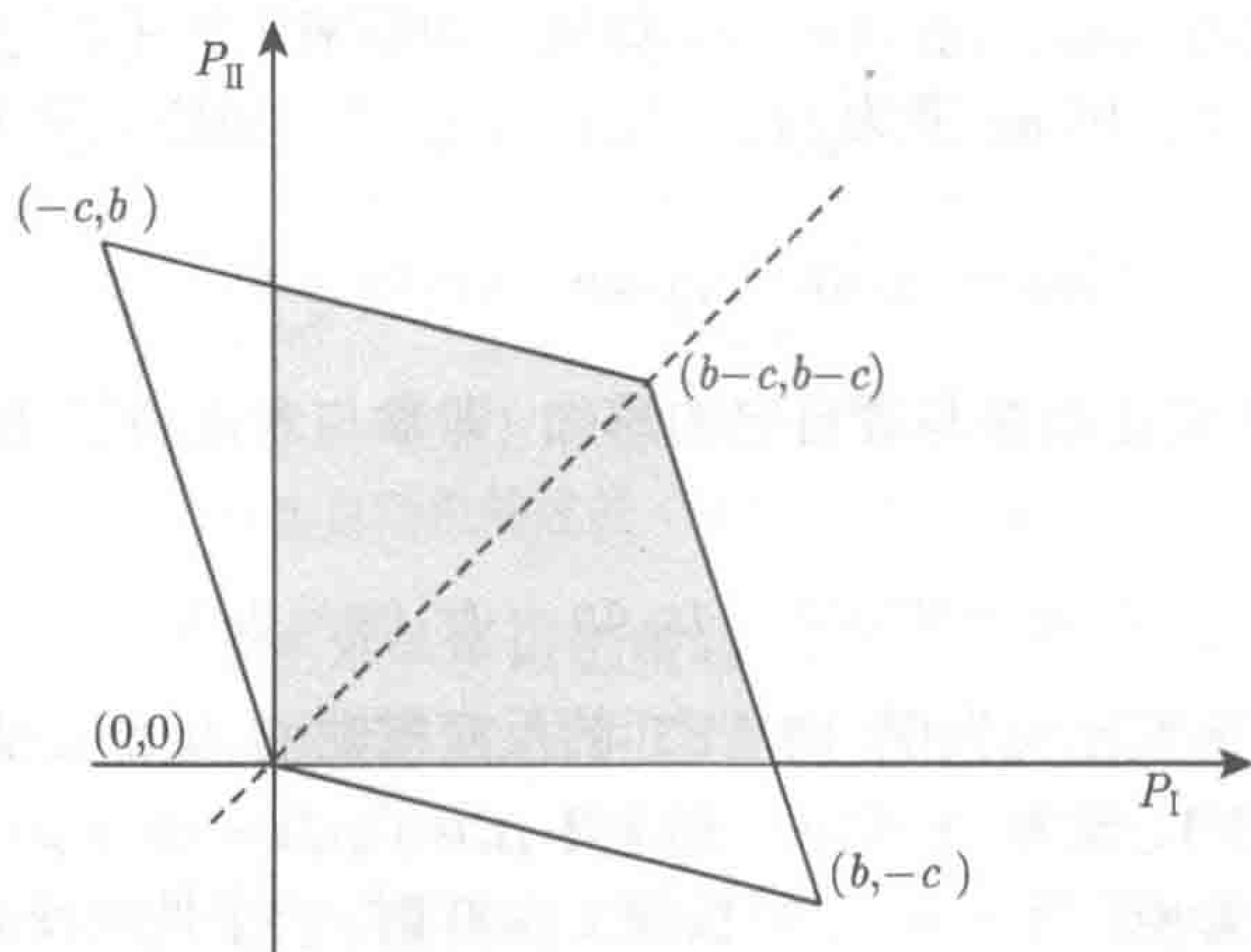


图 3.13　如果参与者 I 和 II 在无限重复的囚徒困境博弈中使用对应的均衡器策略, 则阴影区域中任何收益值对 $(P_{\rm I}, P_{\rm II})$ 都可以达到

这个结果可以沿很多方向推广: 考虑其他重复博弈 (下界 0 替代为最大收益, 即在无视对手的策略时参与者可以保证自己的最大收益), 考虑对未来收益的折扣 (或者允许重复博弈以正的概率停止), 考虑参与者错误实施行为的可能性, 等等.

触发器策略由于其太严厉而经常受到批评: 仅仅因为对手的一次错误 (也许由于不可抗力 (force majeure)), 参与者就将自己置身于永远的毁灭性背叛, 这一点是难以想象的. 尽管如此, 触发器策略对分析两个理性参与者之间的博弈是基本的工具. 然而, 在进化博弈理论中, 触发器策略的作用并不明显.

变体的大众定理可以容易地在单步记忆策略中呈现出来. 事实上, 确实有类似**均衡器** (equalizer) 的策略, 它使得对手总是获得相同的收益而与其策略无关.

在无限重复的囚徒困境中, 对任意 P 与 R 之间的值 π, 存在单步记忆策略 $\boldsymbol{q}=(q_R,q_S,q_T,q_P)$, 使得对手与使用该策略的参与者对局时都获得长期平均收益 π. 相应的反应规则 $\boldsymbol{q}$ 为

$$(1-(R-\pi)a, 1-(T-\pi)a, (\pi-S)a, (\pi-P)a), \tag{3.80}$$

其中 $a>0$ 为满足 $\dfrac{1}{a}\geqslant \max\{T-\pi, R-\pi, \pi-S, \pi-P\}$ 的实数. (a 满足的条件保证了 q_i 是概率.)

事实上, 用 $p_i(n)$ 表示第 n 回合参与者 I 收益为 i 时, 参与者 II 在第 $n+1$ 回合选择合作的条件概率; 并且令 $s_i(n)$ 为收益值取 i 的概率. 由第 n 回合的条件, $s_R(n+1)$ 为

$$\begin{aligned}&s_R(n)p_R(n)[1-(R-\pi)a]+s_S(n)p_S(n)[1-(T-\pi)a]\\&\quad+s_T(n)p_T(n)(\pi-S)a+s_P(n)p_P(n)(\pi-P)a.\end{aligned} \tag{3.81}$$

同样地, $s_S(n+1)$ 为

$$\begin{aligned}&s_R(n)(1-p_R(n))[1-(R-\pi)a]+s_S(n)(1-p_S(n))[1-(T-\pi)a]\\&\quad+s_T(n)(1-p_T(n))(\pi-S)a+s_P(n)(1-p_P(n))(\pi-P)a.\end{aligned} \tag{3.82}$$

将这些表达式相加得到参与者 I 在第 $n+1$ 回合选择合作的概率, 即 $s_R(n+1)+s_S(n+1)$, 其由下式给出:

$$s_R(n)[1-(R-\pi)a]+s_S(n)[1-(T-\pi)a]+s_T(n)(\pi-S)a+s_P(n)(\pi-P)a. \tag{3.83}$$

因此

$$s_R(n)+s_S(n)-s_R(n+1)-s_S(n+1) \tag{3.84}$$

去掉因子 a 为

$$Rs_R(n)+Ss_T(n)+Ts_S(n)+Ps_P(n)-\pi[s_R(n)+s_S(n)+s_T(n)+s_P(n)]. \tag{3.85}$$

由于 $s_i(n)$ 和为 1, 上式恰好为 $A'(n)-\pi$, 其中, $A'(n)$ 是参与者 II 在第 n 回合的收益 (注意, 一个参与者的 S 对应另一个参与者的 T). 对 $n=0,\cdots,N$ 关于上式求和并除以 $N+1$, 有

$$\frac{s_R(0)+s_S(0)-s_R(N)-s_S(N)}{a(N+1)}=\frac{A'(0)+\cdots+A'(N)}{N+1}-\pi, \tag{3.86}$$

从而

$$\lim_{N\to\infty}\frac{A'(0)+\cdots+A'(N)}{N+1}=\pi. \tag{3.87}$$

3.20　交替囚徒困境

在很多现实生活的直接回报实例中, 两个参与者轮流扮演捐赠者和接受者的角色, 然而在大多数文献以及我们目前的讨论中, 总是假设两个参与者都同时扮演捐赠者并做出决定. 通常, 这种假设影响不大. 但在有些情况下, 会产生显著的差异.

在交替的囚徒困境 (或者交替的捐赠博弈) 中, 合作意味着应该合作时选择 C, 这会影响参与者的策略和收益. 例如, 两个 TFT 参与者进行通常的同步重复囚徒困境, 由失误引起的单边背叛, 将导致两个参与者轮流选择背叛. 而在交替囚徒困境中, 如果由于失误导致单边背叛, 结果将是一系列的相互背叛: 两个参与者总是选择 D, 如图 3.14 所示. 事实上, TFT, AllC 和 AllD 两两之间的相互作用都保持不变, 平均收益将与同时决策的情形一样. 但如果两个 Pavlov 参与者相互对局, 结果会非常不同. 一个失误的背叛将由背叛回应, 这将引出合作, 而后又由背叛回应. 因此, 每一个参与者在错误的背叛之后, 会在以后的回合中周期性的出现两次背叛和一次合作. 下一个失误有 2/3 的概率不会影响这个机制. 只以 1/3 的概率将会使博弈恢复到相互合作. 在无限重复的情形 ($w=1$), 平均收益为 $(b-c)/2$.

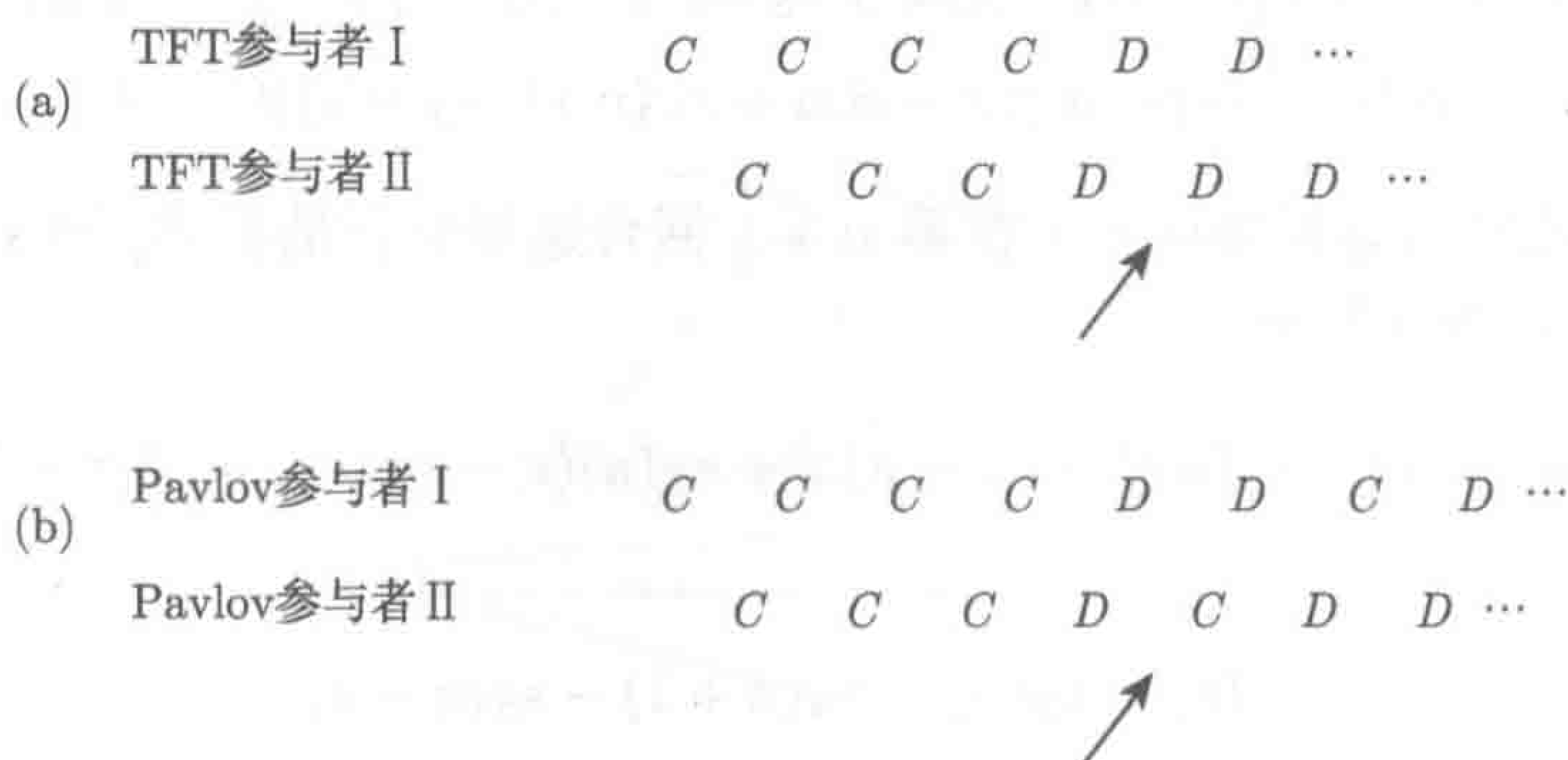


图 3.14　交替囚徒困境中失误导致的背叛所产生的结果. (a) 两个 TFT 参与者之间, (b) 两个 Pavlov 参与者之间. 箭头标出博弈中的失误

表 3.2 具有行为误差的交替的捐赠博弈

*	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
S_0	$\boldsymbol{0}$	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{3}$	b	$\frac{b}{2}$	b	0	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{2}$	b	$\frac{2b}{3}$	b
S_1	$-\frac{c}{2}$	$\boldsymbol{\frac{b-c}{3}}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{3}$	b	b	b	$-\frac{c}{2}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-c}{4}$	b	b	b
S_2	0	$\frac{b-c}{2}$	$\boldsymbol{0}$	$\frac{b-c}{2}$	0	$\frac{b-c}{2}$	0	$\frac{b-c}{2}$	0	$\frac{2b-c}{3}$	0	$\frac{2b-c}{3}$	$\frac{2b-c}{4}$	$\frac{2b-c}{2}$	$\frac{2b-c}{3}$	$\frac{2b-c}{2}$
S_3	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\boldsymbol{\frac{b-c}{2}}$	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$
S_4	$-\frac{c}{3}$	$\frac{b-c}{3}$	0	$\frac{b}{2}$	$\boldsymbol{\frac{b-c}{3}}$	$\frac{2b-c}{3}$	$\frac{2b}{3}$	b	$-\frac{c}{3}$	$\frac{2b-c}{5}$	0	$\frac{b}{2}$	$\frac{3b-c}{6}$	b	$\frac{2b}{3}$	b
S_5	$-c$	$-c$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-2c}{3}$	$\boldsymbol{\frac{b-c}{2}}$	b	b	$-c$	$-c$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-c}{3}$	b	b
S_6	$-\frac{c}{2}$	$-c$	0	$\frac{b-c}{2}$	$-\frac{2c}{3}$	$-c$	$\boldsymbol{\frac{b-c}{2}}$	$\frac{2b-2c}{3}$	$-\frac{c}{3}$	$\frac{b-c}{2}$	0	$\frac{2b-c}{3}$	$\frac{b-c}{2}$	$\frac{4b-3c}{5}$	$\frac{4b-2c}{5}$	$\frac{2b-c}{2}$
S_7	$-c$	$-c$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$-c$	$-c$	$\frac{2b-2c}{3}$	$\boldsymbol{\frac{2b-2c}{3}}$	$-c$	$-c$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-3c}{4}$	$\frac{2b-2c}{3}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$
S_8	0	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{3}$	b	$\frac{b}{3}$	b	$\boldsymbol{0}$	$\frac{3b-c}{5}$	$\frac{b-c}{3}$	$\frac{2b-c}{3}$	$\frac{3b-c}{6}$	$\frac{3b-c}{3}$	$\frac{2b-c}{3}$	$\frac{3b-c}{3}$
S_9	$-\frac{c}{2}$	$\frac{b-c}{3}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$\frac{b-2c}{5}$	b	$\frac{b-c}{2}$	b	$\frac{b-3c}{5}$	$\boldsymbol{\frac{b-c}{2}}$	$b-c$	$b-c$	$\frac{b-c}{2}$	$\frac{3b-c}{3}$	$\frac{3b-2c}{3}$	$\frac{2b-c}{2}$
S_{10}	0	$\frac{b-c}{2}$	0	$\frac{b-c}{2}$	0	$\frac{b-c}{2}$	0	$\frac{b-c}{2}$	$\frac{b-c}{3}$	$b-c$	$\boldsymbol{\frac{b-c}{2}}$	$b-c$	$\frac{b-c}{2}$	$b-c$	$\frac{2b-2c}{3}$	$b-c$
S_{11}	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$\frac{b-2c}{3}$	$b-c$	$b-c$	$\boldsymbol{b-c}$	$\frac{2b-3c}{4}$	$b-c$	$b-c$	$b-c$
S_{12}	$-\frac{c}{2}$	$\frac{b-2c}{4}$	$\frac{b-2c}{4}$	$\frac{b-c}{2}$	$\frac{b-3c}{6}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{3b-2c}{4}$	$\frac{b-3c}{6}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{3b-2c}{4}$	$\boldsymbol{\frac{b-c}{2}}$	$\frac{5b-3c}{6}$	$\frac{5b-3c}{6}$	$\frac{2b-c}{2}$
S_{13}	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$-c$	$\frac{b-2c}{3}$	$\frac{3b-4c}{5}$	$\frac{2b-2c}{3}$	$\frac{b-3c}{3}$	$\frac{b-3c}{3}$	$b-c$	$b-c$	$\frac{3b-5c}{6}$	$\boldsymbol{\frac{2b-2c}{3}}$	$\frac{3b-2c}{3}$	$\frac{3b-2c}{3}$
S_{14}	$-\frac{2c}{3}$	$-c$	$\frac{b-2c}{3}$	$\frac{b-2c}{2}$	$-\frac{2c}{3}$	$-c$	$\frac{2b-4c}{5}$	$\frac{b-2c}{2}$	$\frac{b-2c}{3}$	$\frac{2b-3c}{3}$	$\frac{2b-2c}{3}$	$b-c$	$\frac{3b-5c}{6}$	$\frac{2b-3c}{3}$	$\boldsymbol{b-c}$	$b-c$
S_{15}	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$\frac{b-3c}{3}$	$\frac{b-2c}{3}$	$b-c$	$b-c$	$\frac{b-2c}{2}$	$\frac{2b-3c}{3}$	$b-c$	$\boldsymbol{b-c}$

如果参与者轮流成为潜在的捐赠者, 连续两回合的交替博弈对应于一回合的同步博弈. 假设每一个参与者保持了前两回合的记忆 (即每一个参与者捐赠与否的决定). 博弈结果用 R, S, T, P 表示, 无限重复交替博弈策略由获得 R, S, T, P 收益之后的合作倾向 q_R, q_S, q_T, q_P 来表示. (q_R, q_S, q_T, q_P) 参与者与 (q'_R, q'_S, q'_T, q'_P) 参与者对局的转移概率由如下矩阵给出:

$$\boldsymbol{Q} = \begin{pmatrix} q_R q'_R & q_R(1-q'_R) & (1-q_R)q'_S & (1-q_R)(1-q'_S) \\ q_S q'_T & q_S(1-q'_T) & (1-q_S)q'_P & (1-q_S)(1-q'_P) \\ q_T q'_R & q_T(1-q'_R) & (1-q_T)q'_S & (1-q_T)(1-q'_S) \\ q_P q'_T & q_P(1-q'_T) & (1-q_P)q'_P & (1-q_P)(1-q'_P) \end{pmatrix}, \tag{3.88}$$

这个矩阵与矩阵 (3.67) 完全不同. 收益可以如前计算. 在交替囚徒困境中, Pavlov 失去了很多优势. 如表 3.2 所示, 它会被 "坚定公平" 取代, 后者的反应规则为 $S_{11} = (1, 0, 1, 1)$. 这个策略具有纠错功能, 并且与 S_{14} 和 S_{15} 一样, 在与自己对局时获得的收益最高, 为 $b - c$. 但与 S_{14} 和 S_{15} 相比, 只要 $b > 2c$, "坚定公平" 就不会被其他策略侵入, 如 AllD. 而策略 S_{14} 能通过中性漂变被其他策略侵入. 另一方面, AllD $= S_0$ 总能被 S_8 和 S_{10} 侵入, 后两个又能被 S_{11} 侵入. 如果只考虑实施合作行为中的失误, 如 3.14 节, AllD 能被许多策略通过中性漂变破坏, 这些策略中, S_2, S_6, S_{10} 和 S_{14} 将让位于 "坚定公平".

如果只关注反应策略, 此时 $q_R = q_T = p$ 且 $q_S = q_P = q$, 则捐赠博弈的收益正好与同步博弈一样, 即使行为序列可能完全不同 (如两个 TFT 参与者的博弈). 再一次, GTFT 以胜利者的姿态出现. 在由有限自动机给出的策略中, 在交替和同步的囚徒困境中, CTFT 的收益一样好, 而纠错能力一样弱.

3.21 参考文献

Trivers (1971) 将具有回报的利他主义作为主要因素在合作进化中引入. 重复囚徒困境的研究始于 Axelrod 和 Hamilton (1981), 并由此引发了大量的研究, 见 Axlelrod (1984), Axelrod 和 Dion (1988), Trivers (2006), Kendall, Yao 和 Chong (2007). 之前也有许多理论研究和实证调查, 见 Rapoport 和 Chammah (1965). 关于 *TFT* 在非人类参与者之间的重要性的各种研究, 参阅 Dugatkin (1997), Milinski (1987) 和 Hamerstein (2003). 懦夫博弈, 也称为鹰—鸽博弈, 对进化博弈论的早期发展起到了重要作用, 见 Maynard Smith (1982). Sugden (1986) 强调了雪堆博弈与合作的相关性. 3.2 至 3.4 节关于重复捐赠博弈的论述取自于 Brandt 和 Sigmund (2006), 3.5 至 3.8 节关于反应策略的论述取自于 Nowak 和 Sigmund (1990). 适应动力学由 Nowak 和 Sigmund (1989) 及 Hofbauer 和 Sigmund (1990) 引入, 现在已

经有了很大的发展, 见 Dieckmann 和 Law (1996) 或是 Dieckmann 和 Metz (2009). Molander (1985) 引入了大度的 *TFT*, 也可参阅 Nowak 和 Sigmund (1992). Lindgren (1991) 及 Nowak, Sigmund 和 El-Sedy (1995) 讨论了单步记忆策略, 见 Hilbe (2008) 的详细论述. Pavlov 策略的成功第一次由 Kraines 和 Kraines (1989)、Fudenberg 和 Maskin (1990) 注意到, 其在进化动力学的作用由 Nowak 和 Sigmund (1993) 说明, 其在有限群体中的作用由 Imhof, Fudenberg 和 Nowak (2007) 说明. May (1987), Boyd (1989), Bendor 和 Swistak (1995), Sherratt[①] 和 Roberts (2001) 强调了稳定合作中过失与偶尔背叛的作用. 由有限自动机实施的策略, 见 Aumann (1981), Rubinstein (1986), Abreu 和 Rubinstein (1988), Banks 和 Sundaram (1990), Binmore 和 Samuelson (1992) 以及 Leimar (1997). Boyd (1989) 强调了过失的作用, 也可参阅 May (1987). CTFT 由 Sugden (1986) 引入, 也可参阅 Boerlijst, Nowak 和 Sigmund (1997a). 对于经济学实验, 参阅 Milinski 和 Wedekind (1998), Kollock (1993) 以及 Kagel 和 Roth (1995). 在重复博弈大众理论的大量文献中, 我们特别推荐 Selten 和 Hammerstein (1984) 以及 Fudenberg 和 Maskin (1986). 关于均衡器策略的章节, 参阅 Boerlijst, Nowak 和 Sigmund (1997b). 交替囚徒困境在 Nowak 和 Sigmund (1994), Frean (1994) 以及 Neill (2001) 中有讨论. 关于持续投资水平这一重要主题, 参阅 Roberts 和 Sherratt (1998), Wahl 和 Nowak (1999) 以及 Killingback 和 Doebeli (2002). Fishman, Lotem 和 Stone (2001), Johnson, Stopka 和 Bell (2002) 以及 McNamara, Barta 和 Houston (2004) 强调了行为变化的重要性.

① 译注: 原文此处为 "Sheratt".

第 4 章　间接回报：声誉的作用

4.1　间接回报

前面的章节，我们讨论了与同一对手进行重复捐赠博弈时的策略. 本章，我们假定同样的博弈重复进行，但总是针对不同的对手. 第三方一定会对帮助别人的行为给予回报. 我们将看到，这会使直接回报与间接回报之间产生较大的差异.

在之前的章节中，TFT 策略起到了基本的作用，其特点是根据前一回合的结果来决定本回合的策略行为. 在直接回报时，一个参与者的遭遇是由其对手的策略行为所导致的. 但在间接回报过程中情况则有所不同：两位参与者前一回合的对手与现在的不同. 因此，有两种不同形式的回报. 参与者可以根据其在前一回合的遭遇或者根据其对手在前一回合的策略行为来决定本回合是否捐赠.

简而言之，参与者或者被一种感激之情所影响——“有人帮助了我，我感到愉悦并因此想帮助别人”，或者参与者被一种赞赏之意所支配——“我的对手在与他人博弈时行为高尚，作为回报我将帮助他”.

一种情形，参与者 A 帮助了参与者 B，从而参与者 B 帮助参与者 C. 我们称这类现象为错位 (misguided) 的回报：回报给予了参与者 C 而不是参与者 A. 另一种情形则是参与者 A 帮助了参与者 B，而参与者 C 又来帮助参与者 A，见图 4.1. 这被称为替代 (vicarious) 的回报：A 从第三方 C 而不是从 B 获得了应得的回报. 一种情形，回报者先获得帮助，然后在下一回合中，向并未直接帮助过自己的其他人表达感激. 另一种情形，回报者奖励捐赠者，因为捐赠者曾经对别人付出. 替代的回报和错位的回报也被称为顺向的和逆向的回报. 前一种情形，参与者负担捐赠成本期望将来获得回报，后一种情形，参与者因为此前获得了捐赠而有能力负担.

有趣的是，两种情形看起来都出现在经济学实验中. 但是目前所讨论的理论模型中，替代的回报即赞赏，运行良好，而错位的回报即感激，似乎不易解释.

直接回报与间接回报的另一区别是直接回报的两个参与者经历的博弈回合数相同 (即使他们改变捐赠和接受的角色). 相比之下，间接回报中的两个参与者在历史上仅相遇一次，他们经历的博弈回合数也不一样. 第二个回合的捐赠者可能会与第五回合的接受者配对. 有些模型假设参与者的博弈是同步的，参与者同时开始，以相同的节奏进行博弈，并根据相同的信号进行策略更新. 但这样显得不够自然，因而不予讨论.

对于间接回报而言, 下一回合博弈发生的概率 w 的取值高于直接回报的是不合理的, 因为在现实生活的小群体中, 经历很多回合的参与者很可能会与相同的对手进行多次博弈, 因此他们参与的是直接回报. 但在下面的间接回报模型中, $w = 1$ 的极限情形不改变结果.

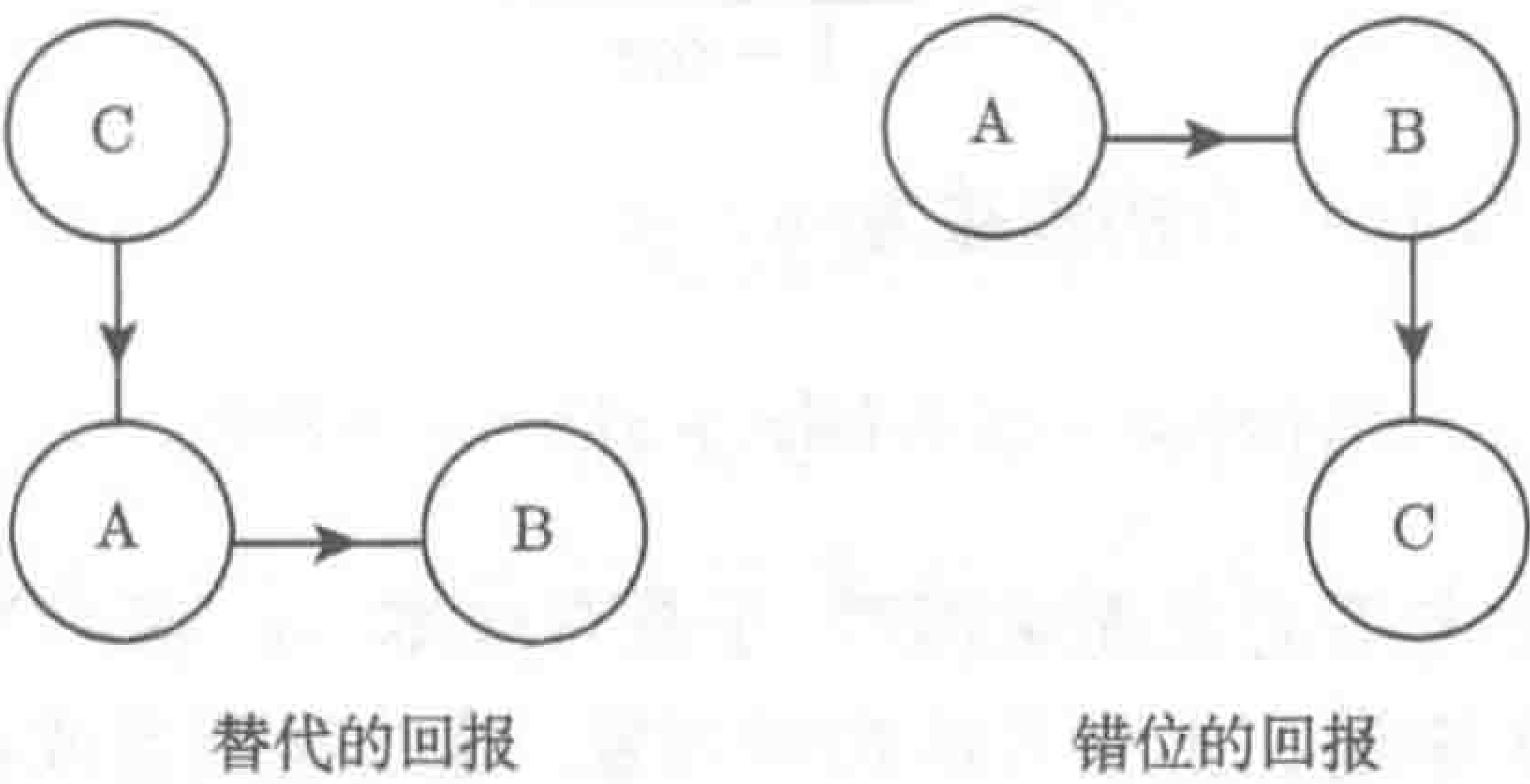


图 4.1 替代的回报和错位的回报. 前者, 参与者 C 观察了参与者 A 和 B 之间的博弈并据此决定是否帮助参与者 A. 另一种情形, 参与者 B 是否转而帮助 C 取决于此前其是否从参与者 A 处获得帮助

4.2 合作者、背叛者以及回报者

我们考虑一个持续进入模型: 参与者依次进入一个大型群体与不同的参与者随机进行异步博弈, 不时更新策略, 并最终退出. 因为群体规模较大, 所以其组成比例变化非常缓慢以至于可以认为它在相邻回合之间保持不变. 可以假设在任一回合中, 随机指定某一方为捐赠者, 另一方为接受者. 为简化起见, 我们假设参与者既作捐赠者又作接受者. 参与者的捐赠方和接收方虽然不同, 但是仍然可以采用捐赠博弈的支付矩阵 (3.1) 来描述每回合捐赠博弈的收益.

考虑替代回报. 分别记无条件策略 AllC, AllD 及回报者的频度为 x, y 和 z. 首先考虑回报者中最简单的一种, 即直接回报中的 TFT. 如果得知当前的对手在前一回合中曾经拒绝捐赠, TFT 会在当前回合拒绝捐赠. 记 q 为参与者知道 (通过旁观或者闲聊) 其随机分配的对手在前一回合中策略行为的概率. 假定在最初没有任何历史博弈信息时, 回报者采取信任原则, 即认为接受者在之前都选择了捐赠.

如同直接回报中容许出错, 记 ϵ 为参与者未能成功实施捐赠意图的概率. 这可能源于失误, 也可能因为外部原因, 譬如暂时的资源缺乏. 假设拒绝的意图总能准确无误地实施.

记 h 为参与者拥有好声誉 (即在前面回合中提供帮助) 的频度. 因为群体数量庞大, 所以频度 h 在相邻回合保持不变, 记 $\bar{\epsilon} = 1 - \epsilon$, 则

$$h = \bar{\epsilon}[x + z(1 - q + qh)]. \tag{4.1}$$

实际上, 愿意捐赠的参与者只能是 AllC 类型参与者 (概率为 x), 或者是不知道对手声誉 (概率为 $1-q$) 的回报者 (概率为 z) 以及知道对手拥有好声誉 (概率为 qh) 的回报者. 愿意捐赠的概率 (方括号中的项) 需要乘以 $\bar{\epsilon}$ (不出错的概率). 因此

$$h = \frac{\bar{\epsilon}(x + z(1-q))}{1 - \bar{\epsilon}qz}. \tag{4.2}$$

AllC 参与者第 n $(n \geqslant 1)$ 回合的收益为

$$P_x(n) = -c\bar{\epsilon} + b\bar{\epsilon}[x + z(1 - q + \bar{\epsilon}q)]. \tag{4.3}$$

实际上, 一个 AllC 参与者总是愿意捐赠, 于是有成本 $-c$ (概率为 $\bar{\epsilon}$). 另一方面, 该参与者同时也是对手中潜在捐赠者的捐赠对象. 这些捐赠者或者是无条件合作者 (概率为 x), 或者是不知道对手声誉 (概率为 $1-q$) 的回报者 (概率为 z), 以及知道这个参与者声誉 (概率为 q) 且其恰为好声誉 (概率为 $\bar{\epsilon}$) 的回报者. 很明显这是因为 AllC 参与者只有在前一回合捐赠出错才会拥有坏声誉. 因为捐赠存在 ϵ 的出错概率, 所以最终获得的收益为 $b\bar{\epsilon}$.

类似的, 背叛者第 n 回合的收益为

$$P_y(n) = b\bar{\epsilon}[x + (1-q)z], \tag{4.4}$$

而对于回报者, 称之为 A, 有

$$P_z(n) = -c\bar{\epsilon}(1 - q + qh) + b\bar{\epsilon}[x + z[1 - q + \bar{\epsilon}q(1 - q + qh)]]. \tag{4.5}$$

和式中的第二项 (去掉期望收益 $b\bar{\epsilon}$) 为对手愿意向参与者 A 捐赠的概率. 这包括或者对手为 AllC(概率为 x), 或者对手为回报者 (概率为 z), 其可能不知道参与者 A 的声誉 (概率为 $1-q$), 也可能知道参与者 A 的声誉 (概率为 q) 且恰好 A 有好声誉. 而参与者 A 要拥有好声誉, 其只有在前一回合愿意捐赠 (或者不知道对手的声誉, 或者知道对手有好声誉 (概率为 h)) 并且成功地实施了捐赠 (概率为 $\bar{\epsilon}$).

直接计算可得

$$P_z(n) - P_y(n) = [P_x(n) - P_y(n)](1 - q + qh). \tag{4.6}$$

虽然初始回合 (标记为 0) 的收益稍有不同: $P_x(0) = -c\bar{\epsilon} + b\bar{\epsilon}(x+z), P_y(0) = b\bar{\epsilon}(x+z), P_z(0) = -c\bar{\epsilon}[1 - q + qh] + b\bar{\epsilon}(x+z)$, 但依然满足此式. 因此 3.2 节给出的每一回合收益 P_x, P_y 以及 P_z 也满足

$$P_z - P_y = [P_x - P_y](1 - q + qh). \tag{4.7}$$

显然 $P_x(n)-P_y(n)=\bar{\epsilon}(-c+b\bar{\epsilon}qz)(n\geqslant 1), P_x(0)-P_y(0)=-c\bar{\epsilon}$. 于是每回合的收益满足

$$P_x-P_y=\bar{\epsilon}(-c+wb\bar{\epsilon}qz). \tag{4.8}$$

当 $w=1$ 时, 等式也成立.

4.3 复制动力学

为了研究复制方程 $\dot{x}=x(P_x-\bar{P})$ 等, 可以将每种类型的收益减去 P_y 并除以 $\bar{\epsilon}$. 对得到的表达式依然采用原记号 P_x, P_y 和 P_z. 显然有

$$P_x=f,\quad P_y=0,\quad P_z=f(1-q+qh), \tag{4.9}$$

其中

$$f=-c+wb\bar{\epsilon}qz. \tag{4.10}$$

首先, 考虑去掉因子 f (即把 f 替换为 1) 所得到的复制方程. h 由 (4.2) 给出, 即分母 $1-\bar{\epsilon}qz$ 恒正的分式. 在方程右端乘以该恒正项不会改变解的轨线 (只是改变了解的速度). 由此得到具有如下收益的复制方程

$$P_x=1-\bar{\epsilon}qz,\quad P_y=0,\quad P_z=1-q+\bar{\epsilon}qx. \tag{4.11}$$

当 $q<1, \epsilon>0$ 时, $0=P_y<P_z<P_x$. 因此, 除在边界 $x=0$ 上, 轨线收敛到 $z=1$ 以外, S_3 内的其他所有轨线都收敛到 $x=1$. 直接计算可得, 存在不变运动 $V=zx^{q-1}y^{-\epsilon q}$.

当 $\epsilon=0$ (没有出错) 时, 边界 $y=0$ 为不动点集, 不变运动为 $V=zx^{q-1}$. 当 $q=1$ (对手的信息完全公开) 时, 边界 $x=0$ 为不动点集, 而不变运动为 $V=zy^{-\epsilon}$. 现在考虑由 (4.9) 给出的 (即含有因子 f) 的复制动力系统.

当 $q<c/wb\bar{\epsilon}$ 时, 如果 z 位于 0 到 1 之间, 则 f 为负, 这在整个单形 S_3 上也成立. 所以乘以因子 f, 相当于对系统作了逆时处理. 这意味着 AllC 参与者被回报者和背叛者共同占优, 而回报者被背叛者占优. 单形内部的所有轨线都由 $x=1$ (仅有 AllC 参与者) 指向 $y=1$ (仅有 AllD 参与者). 因此, 如果参与者了解对手历史的概率 q 很小时 (即声誉传播范围很小), 合作将不会成为进化的方向. 见图 4.2. 如果 $q>c/wb\bar{\epsilon}$, 则直线 $z=c/wbq\bar{\epsilon}$ 穿过 S_3 的内部, 并成为不动点集的线段. 实际上, 在该线上, $0=P_y=P_x=P_z$, 这些不动点都是纳什均衡. 在单形 S_3 上, 当 f 为 1 时, 所有轨线位于相同的曲线上, 使得 $zx^{q-1}y^{-\epsilon q}$ 变为常数运动. 但与前面的情形不同, 轨线方向在 $z>c/wbq\bar{\epsilon}$ 的区域中保持不变, 如图 4.3 所示.

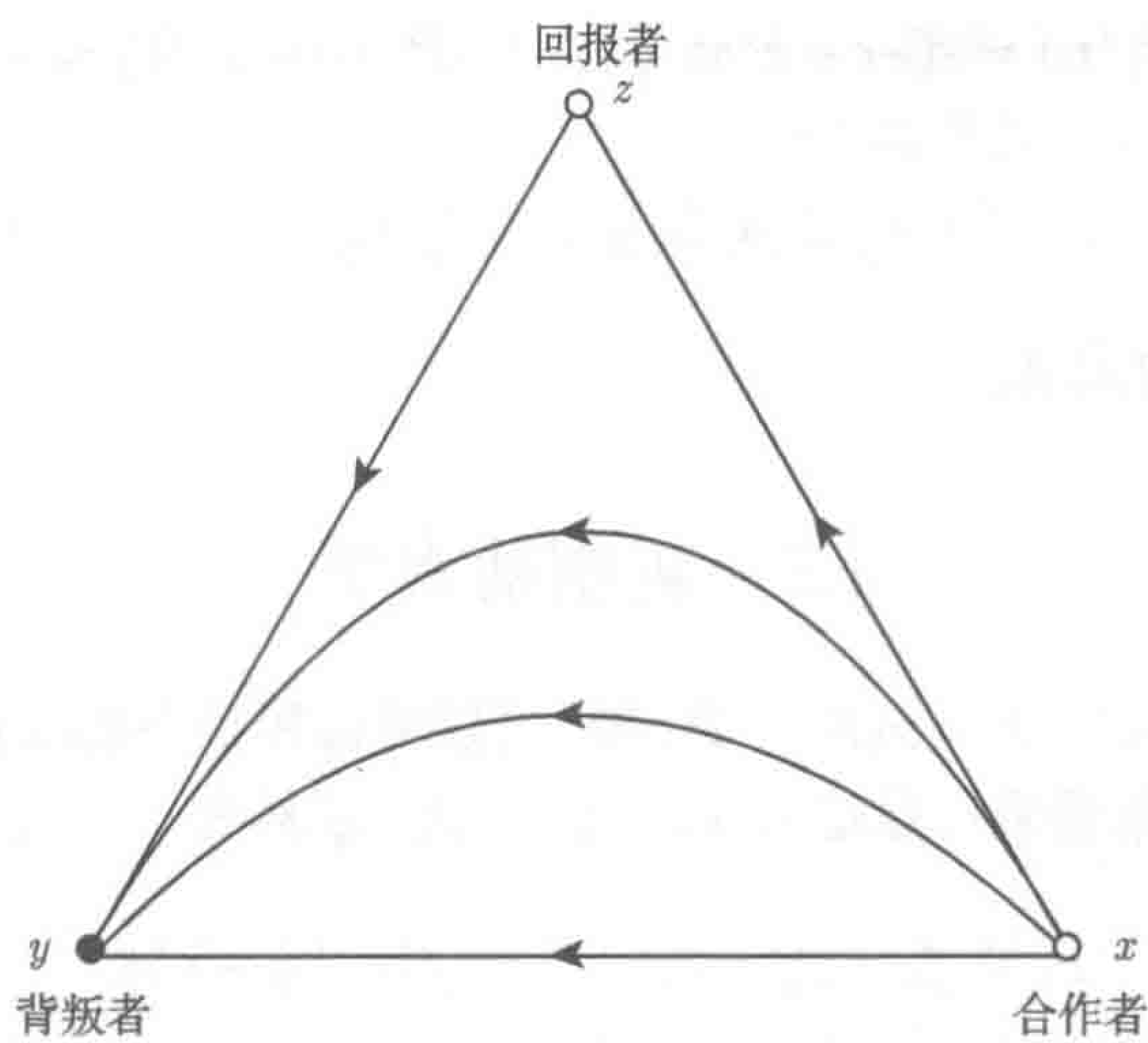


图 4.2　如果参与者不能获得对手足够多的信息, 那么背叛者将胜出

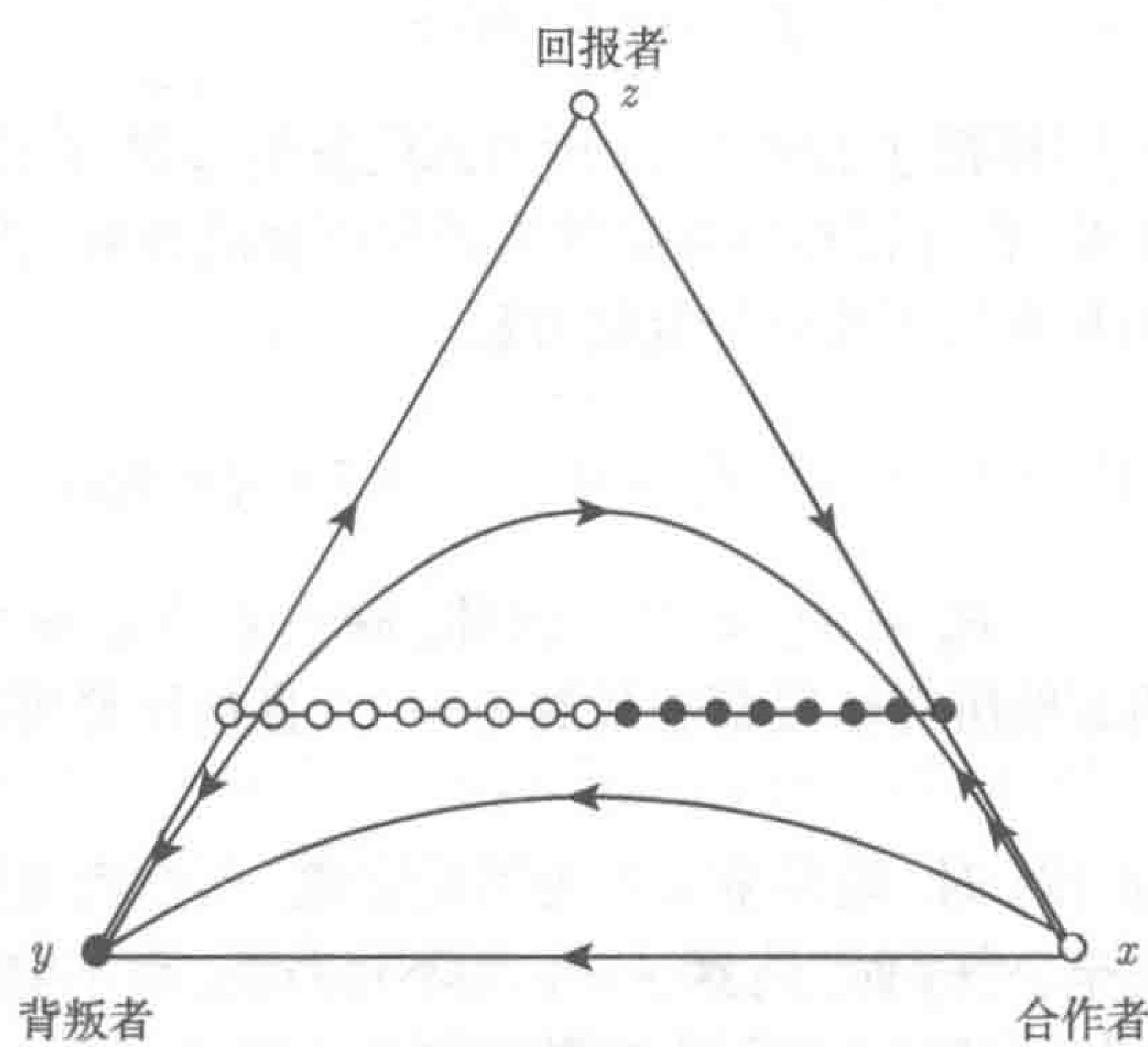


图 4.3　如果信息足够大, 背叛者, 合作者和回报者将共存. 不动点集构成的水平线包含一个 AllC 参与者和回报者的稳定状态. 但长期来看, 随机冲击会导致群体中最终只包含背叛者 (在对参数作貌似合理的取值后, 稳定不动点即实心圆点的部分, 远比图示的部分小)

特别地, 这意味着由 $z = c/wbq\bar{\epsilon}$ 以及 $y = 0$ 所给定的 AllC 参与者和回报者的混合状态对应于该复制动力系统的不动点. 如果群体内的信息公开程度足够高时, 这两类利他主义者 (一类为有条件利他, 另一类为无条件利他) 构成的合作群体将会出现. 这一均衡是稳定的. 然而由于它属于一条不动点集的线段, 因此它不是渐近稳定的.

水平线段的纳什均衡其邻域的动力学行为非常有趣. 线段的一部分是横截稳

定的 (transversely stable), 即动力学特性将抵消微小扰动所导致的离开线段的趋势. 而另一段上扰动则会被动力学特性放大. 朝向高于均衡点线段 z 值方向的微小偏离, 将导致回报者数量先增加然后减少, 最后回到均衡点线段上的稳定部分. 相比之下, 不动点集线段的不稳定部分朝向低于均衡点线段 z 值的微小偏离, 将最终趋于仅有背叛者存在的顶点.

在 $\epsilon = 0$ (没有失误) 的极限情形, 边界 $y = 0$ 由不动点构成, 其中满足 $z \geqslant c/wbq$ 的部分为纳什均衡. 直线 $z = c/wbq$ 也由不动点构成. 低于此直线的所有轨线收敛到 $y = 1$. 高于此直线的每条轨线都收敛到 $y = 0$ 上的某一个纳什均衡, 如图 4.4 所示. 在极限情形 $q = 1$ (信息完全), 边界 $x = 0$ 由不动点构成, 其中 $z < c/wb\bar{\epsilon}$ 部分为纳什均衡. 直线 $z = c/wbq$ 由稳定的不动点集合构成. 此时, 动力学行为如图 4.5 所示.

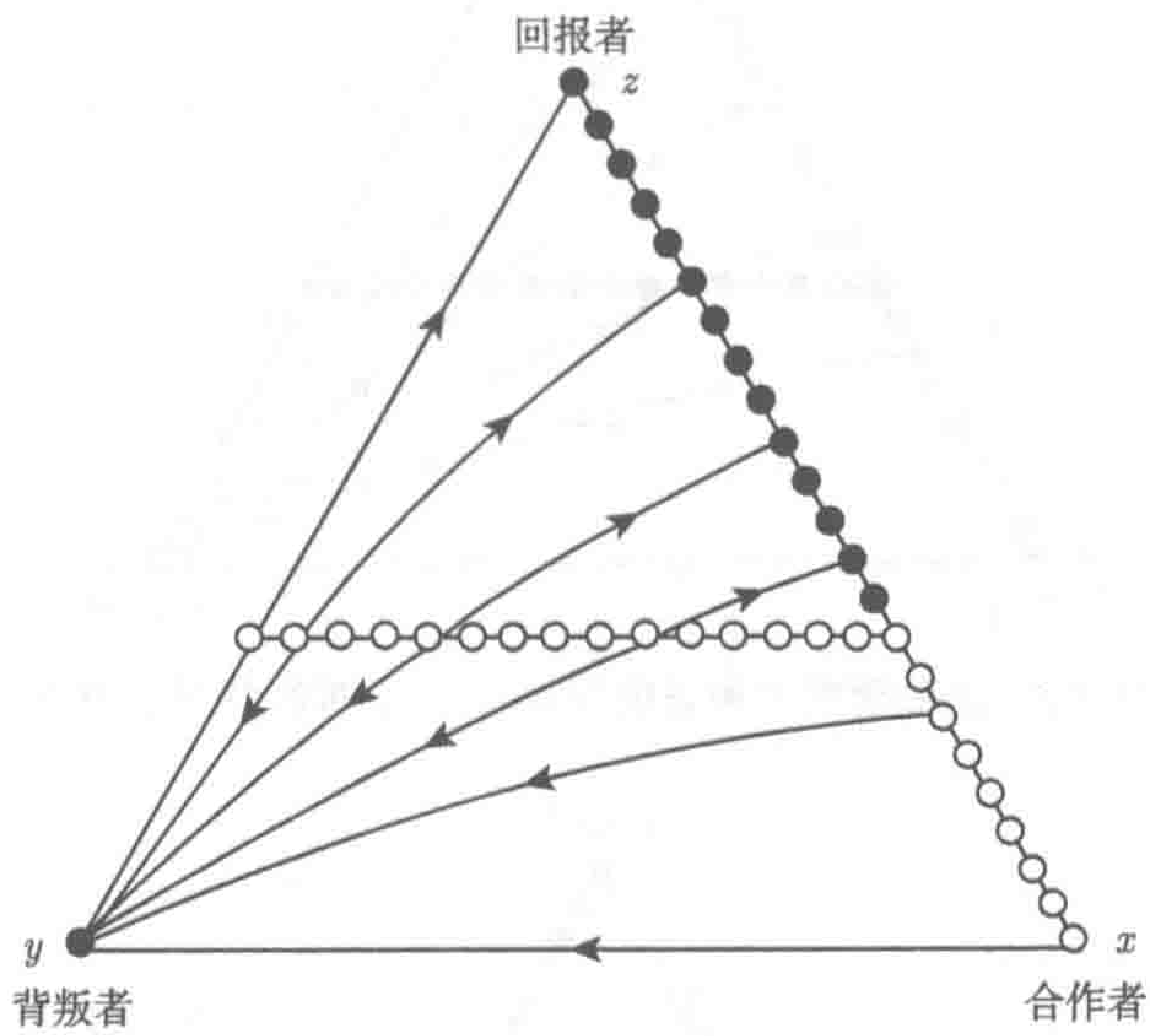


图 4.4 当失误率 ϵ 为 0 时的复制动力学

当 $q = 1$ 和 $\epsilon = 0$ 同时成立时, 边界 $x = 0$ 和 $y = 0$ 都由不动点构成. 在 S_3 的内部, 所有轨线都平行于边界 $z = 0$. 当 $z > c/wb$ 时, 轨线从左向右 (背叛者将消亡), 而当 $z < c/wb$ 时, 轨线从右向左 (无条件合作者将消亡). 直线 $z = c/wb$ 仍然是由纳什均衡构成. 动力学行为如图 4.6 所示.

最后, 我们对于错位的回报作简要讨论. 此时, 回报者依据自己在前一回合中获得帮助与否来决定本回合是否给予帮助. 假设 h 为参与者在前一回合获得帮助的概率, 即有 $h = \bar{\epsilon}(x + hz)$.

第 n 回合 AllC 参与者, AllD 参与者以及回报者的收益分别为 $P_x(n) = -c\bar{\epsilon}+hb$, $P_y(n) = hb$ 和 $P_z(n) = -ch\bar{\epsilon}+hb$. 假设回报者在初始回合总是选择合作, 则 $P_z(0) =$

$-c\bar{\epsilon}+bh$. 同样对总收益进行规范化处理使得 $P_y=0$, 去掉因子 $(1-w)^{-1}$ 可得

$$P_x=-c\bar{\epsilon}, P_z=P_x[1-w(1-h)]. \tag{4.12}$$

此时, 动力学行为和图 4.2 类似: 背叛者顶点 $y=1$ 全局吸引. 即便修改出错概率或假设回报者初始回合背叛, 此结论依然成立. 特别地, 当 $\epsilon\to 0$ 或 $w\to 1$ 时, 结果不变. 令人惊讶的是, 一些实验表明 (实际上也可以想象得到) 基于错位的感激所产生的间接回报并不罕见.

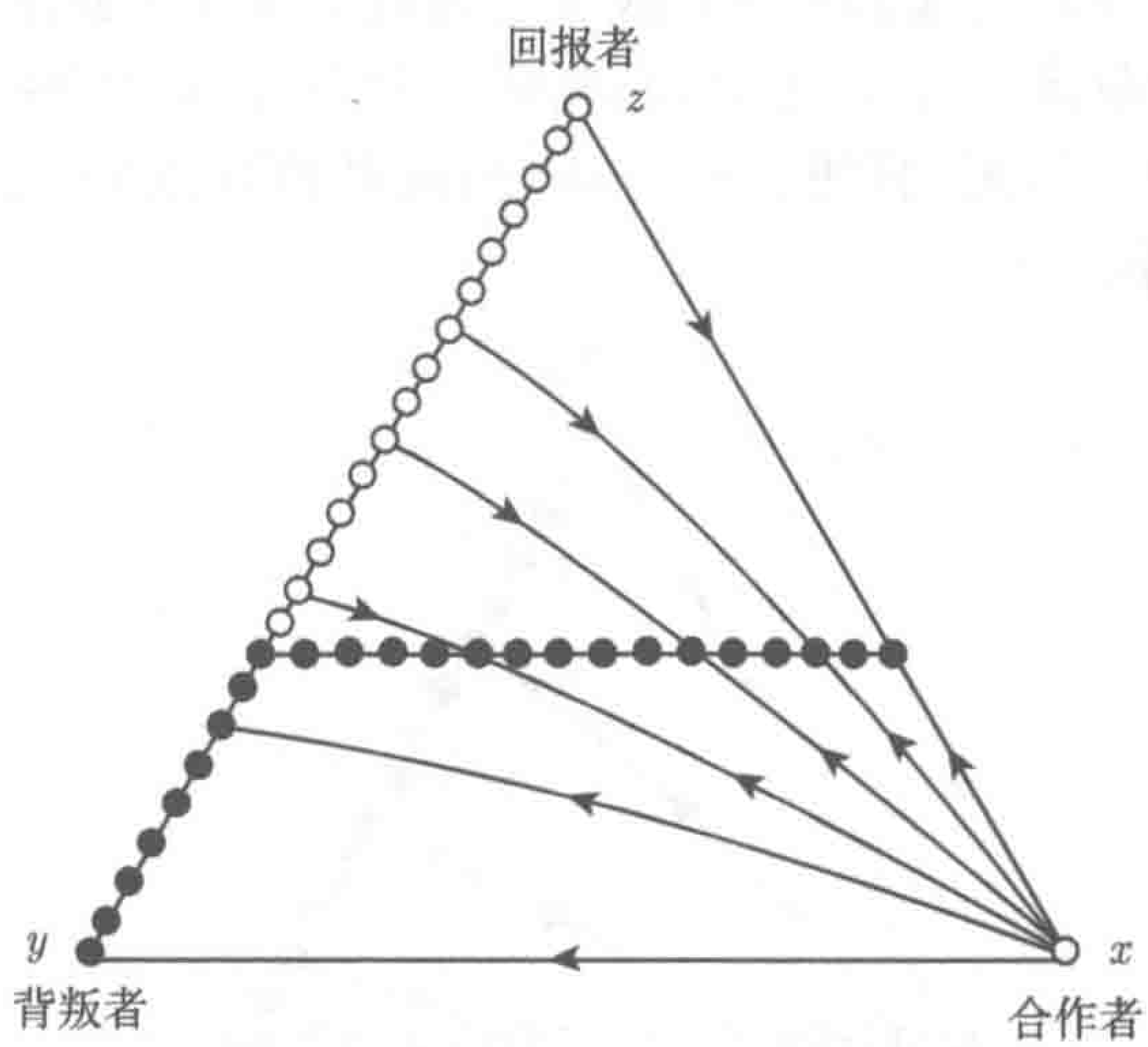


图 4.5 当知道对手声誉的概率 q 为 1 时的复制动力学

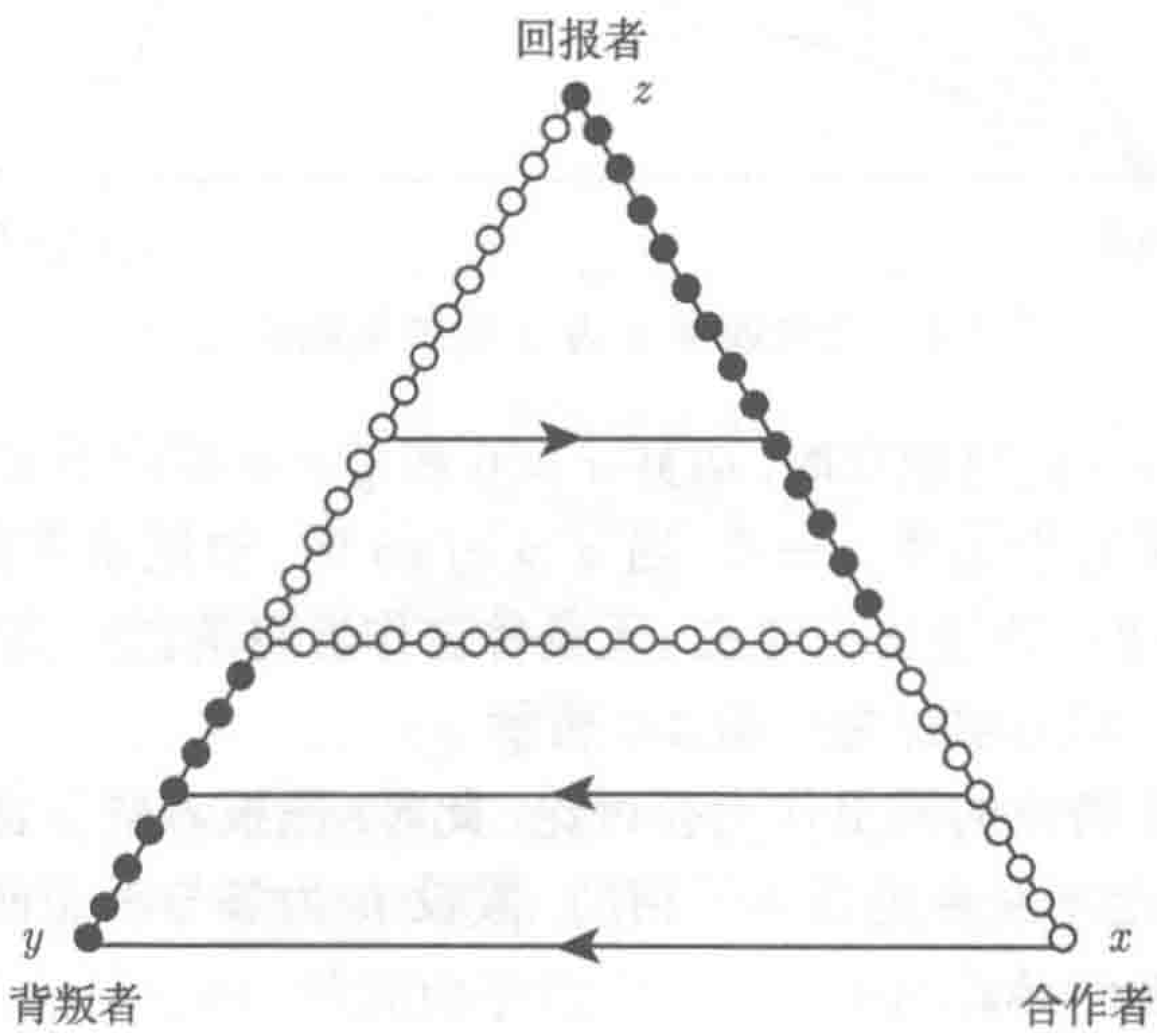


图 4.6 当 $q=1$ 且 $\epsilon=0$ 时的复制动力学

4.4 信任还是怀疑

到目前为止, 我们总是假设回报者是信任的, 即他们总是倾向于认为其对手是值得帮助的. 但如果不清楚对方的声誉, 则如何应对这些对手就是个问题. 因此, 我们引入多疑的回报者, 即在不清楚对手的历史表现时拒绝给予帮助.

假设除了无条件利他者和背叛者 (频度分别为 x 和 y) 之外, 还有信任的和多疑的回报者 (频度分别为 z 和 ζ). 同样允许出错, 即捐赠未成功实施 (而拒绝总是成功实施), 概率为 ϵ. 将 (4.13) 到 (4.17) 中的捐赠意图的概率乘以 $\bar{\epsilon}=1-\epsilon$.

具有好声誉的参与者的频度 (前面回合已经表现) 满足

$$h=\bar{\epsilon}[x+\zeta qh+z(1-q-qh)]. \tag{4.13}$$

($1-q-qh$ 为参与者 z 在前一回合赢得好声誉的概率.) 第 $n-1$ ($n\geqslant 1$) 回合各类参与者的收益为

$$P_x(n)=-c\bar{\epsilon}+b\bar{\epsilon}[x+z(1-q+\bar{\epsilon}q)+\zeta q\bar{\epsilon}], \tag{4.14}$$

$$P_y(n)=b\bar{\epsilon}[x+(1-q)z], \tag{4.15}$$

$$P_z(n)=-c\bar{\epsilon}(1-q-qh)+b\bar{\epsilon}[x+z(1-q+\bar{\epsilon}q(1-q+qh))+\zeta q\bar{\epsilon}(1-q+qh)], \tag{4.16}$$

$$P_\zeta(n)=-c\bar{\epsilon}qh+b\bar{\epsilon}[x+z(1-q+\bar{\epsilon}qh)+\zeta q\bar{\epsilon}qh]. \tag{4.17}$$

因此,

$$P_x(n)-P_y(n)=-c\bar{\epsilon}+b\bar{\epsilon}^2q(z+\zeta), \tag{4.18}$$

$$P_z(n)-P_y(n)=(1-q+qh)[P_x(n)-P_y(n)], \tag{4.19}$$

且

$$P_\zeta(n)-P_y(n)=qh[P_x(n)-P_y(n)]. \tag{4.20}$$

易知对于初始回合 (当 $n=0$ 时), 后三个关系式也成立. 于是可得每一回合的总收益为

$$P_z-P_y=(P_x-P_y)(1-q+qh), \tag{4.21}$$

和

$$P_\zeta-P_y=(P_x-P_y)qh. \tag{4.22}$$

并且有

$$P_x-P_y=\bar{\epsilon}[-c+bwq\bar{\epsilon}(z+\zeta)]:=f. \tag{4.23}$$

考虑复制方程 $\dot{x} = x(P_x - \bar{P})$ 等. 将所有收益项都减去 P_y, 并首先考虑公共因子 f 取 1 时的方程. 于是有

$$P_x = 1, \quad P_y = 0, \quad P_z = 1 - q + qh, \quad P_\zeta = qh. \tag{4.24}$$

由方程 (4.13), 可得

$$h = \frac{\bar{\epsilon}[x + z(1-q)]}{1 - q\bar{\epsilon}(z+\zeta)}. \tag{4.25}$$

如果对 (4.24) 中所有方程右端乘以 h 的分母, 即正函数 $1 - q\bar{\epsilon}(z+\zeta)$, 则所得复制方程与原方程具有相同轨线. 这样, 复制方程满足: $P_y = 0$,

$$P_x = 1 - q\bar{\epsilon}(z+\zeta), \tag{4.26}$$

$$P_\zeta = q\bar{\epsilon}[x + z(1-q)], \tag{4.27}$$

以及

$$P_z = (1-q)[1 - q\bar{\epsilon}(z+\zeta)] + q\bar{\epsilon}[x + z(1-q)]. \tag{4.28}$$

注意到

$$P_z = P_\zeta + (1-q)P_x = (1-q)(1-q\bar{\epsilon}\zeta) + q\bar{\epsilon}x. \tag{4.29}$$

由平均收益 $\bar{P} = xP_x + yP_y + zP_z + \zeta P_\zeta$, 可得

$$\bar{P} = x + z(1-q). \tag{4.30}$$

因此, $P_\zeta = q\bar{\epsilon}\bar{P}$. 进而得

$$P_x - P_z = q[1 - \bar{\epsilon}(q\zeta + x + z)]. \tag{4.31}$$

由于 $0 = P_y < P_\zeta < P_z < P_x$, S_4 内部没有不动点. 在 S_4 的边界上, 轨线走向见图 4.7: 特别地, 边界 $x = z = 0$ 由不动点构成. 当 $\epsilon = 0$ 时, 边界 $y = \zeta = 0$ 也由不动点构成. 函数 $V = (y/x)^{1-q}(z/\zeta)$ 是一个不变运动.

如果考虑方程 (4.21)~(4.23) (即带有因子 f) 的完整的复制动力学行为, 可以发现平面 $z + \zeta = c/bwq\bar{\epsilon}$ 由不动点集构成. f 在此集合上为零. 当 $q > c/wb\bar{\epsilon}$ 时, 此平面与单形相交. 在棱柱 $z + \zeta > c/bwq\bar{\epsilon}$ 内, 轨线方向保持不变, 而在与其互补的多面体内, 轨线方向相反. 此平面内存在稳定不动点集: 这些点为四策略共存点, 如图 4.8 所示. 如果回报者足够多, 信任是值得的.

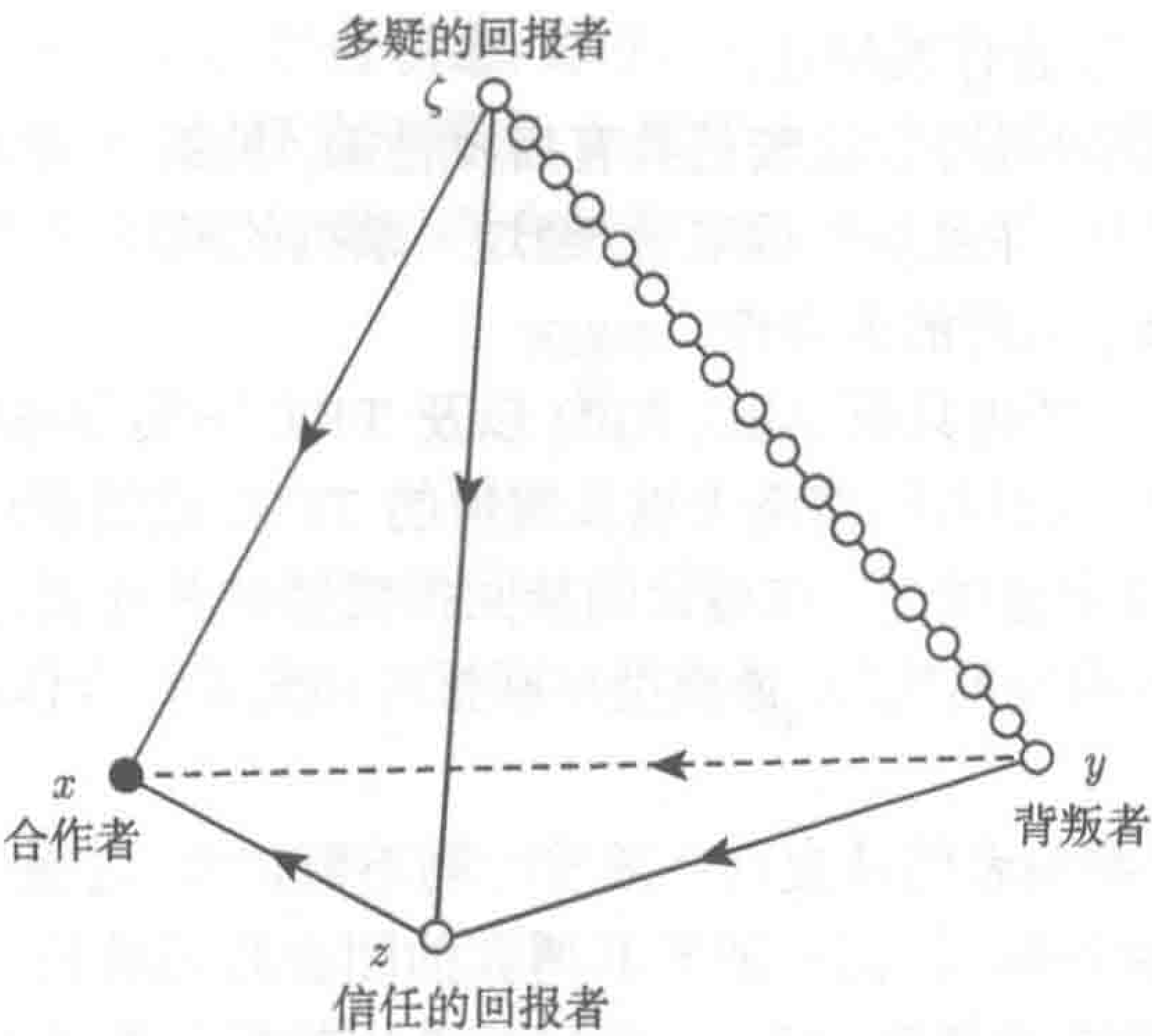


图 4.7 由方程 (4.24) 和 (4.25) 的复制动力学行为

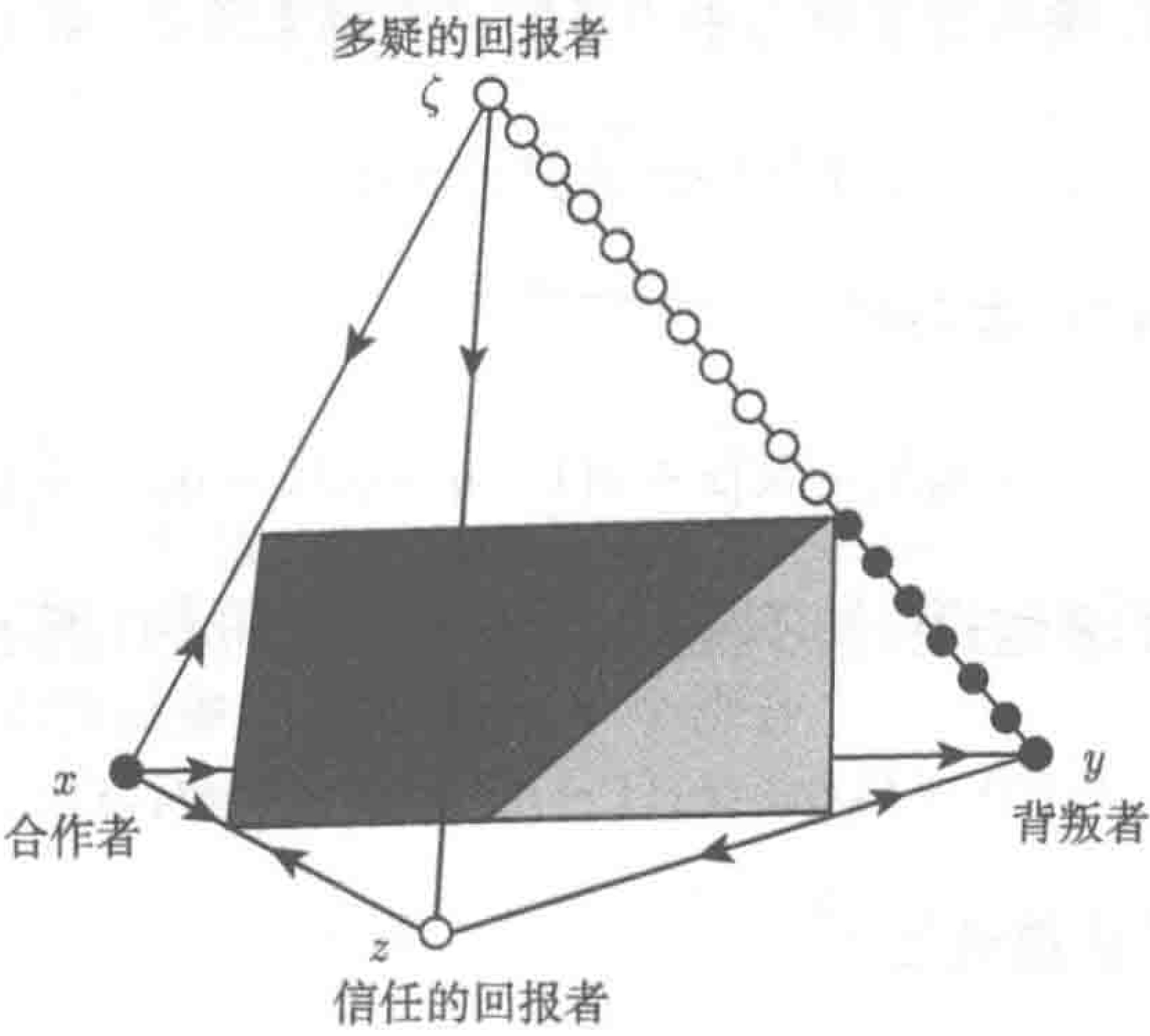


图 4.8 当平均信息分享度 q 足够大时, 合作者, 背叛者, 信任和多疑的回报者的复制动力学行为. 平面集由稳定和不稳定的不动点构成 (对应于深色和浅色区域)

4.5 增长的认知

仅有回报者的群体是不稳定的: AllC 参与者可以成功侵入, 最后会达到 AllC 参与者和回报者的混合均衡. 如图 4.3 所示, 此均衡点稳定, 但非渐近稳定. 该均衡点位于 S_3 内不动点集线段的末端. 容易想象, 小的随机扰动将使群体状态沿此线段移至不稳定驻点的区域. 然后再经过小扰动可能又回到稳定状态, 也可能进入

AllD 的吸引盒, 而导致合作的终止.

将此情形与直接回报进行比较是具有启发性的, 见第 3 章图 3.2 所示. 这里, 也存在 AllD 的吸引盒, 并且从长远来看, 经过一系列的随机扰动, 最终任何群体都会被吸引到这个区域, 从而消灭合作.

在直接回报情形, 考虑只有 AllC, AllD 以及 TFT 三类参与者的简单模型只是第一步. 进一步地, 可以引入巴甫洛夫以及悔悟的 TFT, 这将展现更好的合作前景. 这些结论在间接回报中也成立. 在讨论间接回报模型的其他更为复杂的策略之前, 我们先研究一个初始模型的变种, 该模型将回报者和无条件合作者的驻点变成了稳定的吸引子.

现在假设, 每个参与者的社交网络随着时间不断扩大. 这意味着参与者知道其对手声誉的概率不再是常量, 而是随着其博弈的回合数而增长. 假设 q_n 为回合 n 时参与者了解对手声誉的概率. 记 w_n 为随机选中的参与者处于第 n 回合的概率 (初始回合为 $n=0$). 记 q 为 q_n 的平均 (关于整个群体), 即 $q=\sum w_n q_n$. 如果对所有 n, $q_n > q_{n-1}$, 即参与者了解对手声誉的程度越来越高. 取 $q_{-1}=0$, 于是有

$$q > s := \sum w_n q_{n-1}. \tag{4.32}$$

易见, 对应于方程 (4.5), 我们有

$$P_z(n) = -c\bar{\epsilon}(1-q_n+q_n h) + b\bar{\epsilon}[x + z(1-q+q\bar{\epsilon}(1-q_{n-1}+q_{n-1}h))], \tag{4.33}$$

而 $P_x(n)$ 和 $P_y(n)$ 的表达式保持不变, 如方程 (4.3) 和 (4.4) 所示. 由此得到

$$P_z(n) - P_x(n) = \bar{\epsilon}(1-h)(cq_n - zb\bar{\epsilon}qq_{n-1}) \tag{4.34}$$

记 $z_{cr} := c/b\bar{\epsilon}s$, 则总收益满足

$$P_z(z_{cr}) = P_x(z_{cr}). \tag{4.35}$$

假设 $z_{cr} < 1$, 此条件要比之前的条件 $q > c/b\epsilon$ 稍强. 因为方程 (4.3) 和 (4.4) 蕴含了当 $n>0$ 时, $P_x(n) - P_y(n) = -c\bar{\epsilon} + b\bar{\epsilon}^2 qz$. 这样, 当 $z = z_{cr}$ 时, 表达式

$$P_x(n) - P_y(n) = c\bar{\epsilon}(q-s)/s > 0 \tag{4.36}$$

对于 $n>0$ 恒成立. 对于初始回合, 有 $P_x(0) - P_y(0) = -c\bar{\epsilon}$, 其为负. 这样, 如果 w_0 足够小, 即足够多的参与者都经历了多个回合的博弈, 那么

$$P_x(z_{cr}) > P_y(z_{cr}). \tag{4.37}$$

因此, 存在 AllC 和回报者的混合状态, 即 $F_{xz} = (1 - z_{cr}, 0, z_{cr})$. 在此状态, AllD 无法成功侵入. 复制动力学最终为双稳态: 一个吸引子为全部是背叛者的状态, 另一个是回报者与无条件利他主义者的混合状态, 如图 4.9 所示.

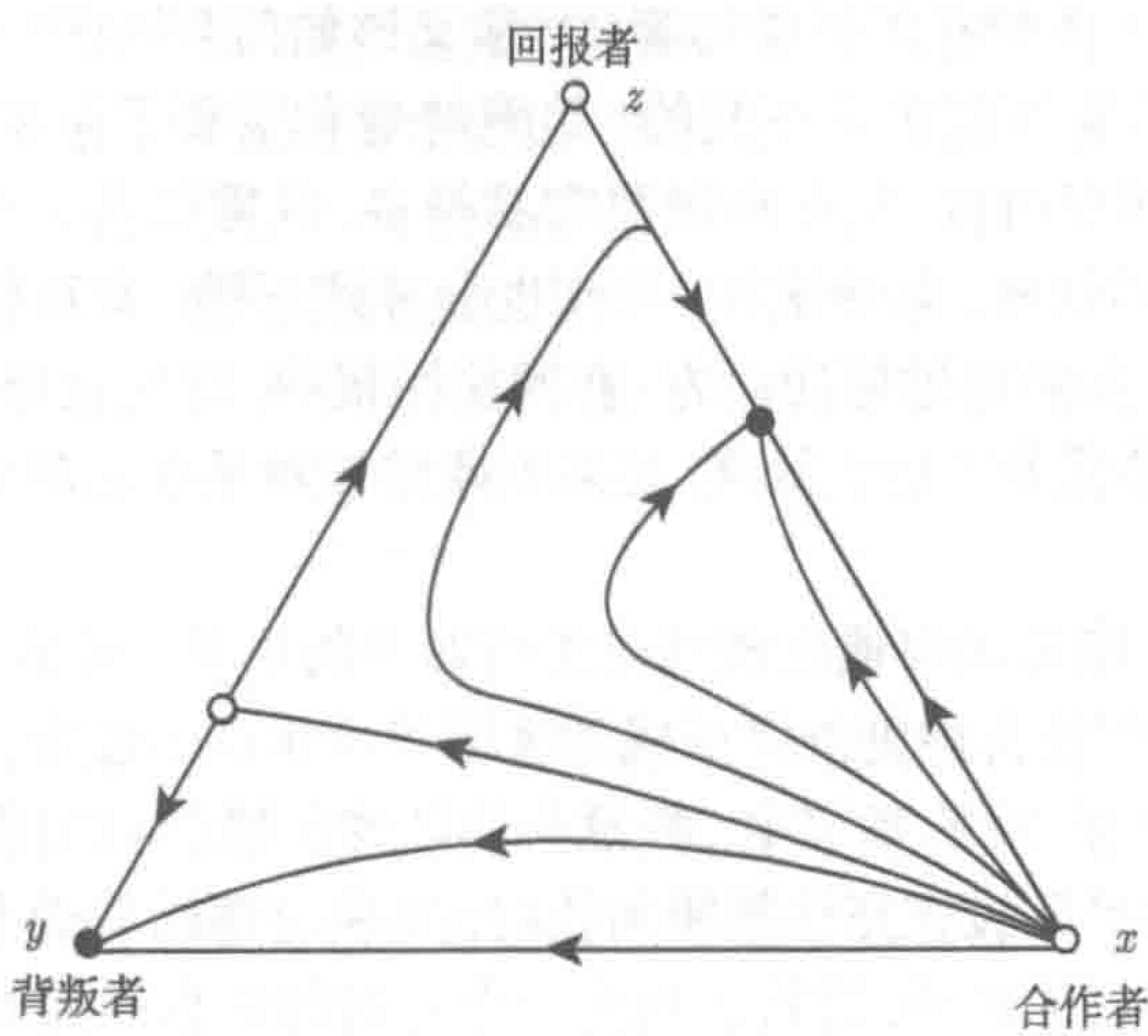


图 4.9 当参与者的社交网络随着时间而扩大 (且足够多的参与者经历多个回合的博弈) 时的复制动力学

这一结果是有瑕疵的, 因为可能我们放弃了通常的假设: w_n 正比于 w^n (3.2 节). 事实上, 在那种情形, $(q - s)/s = (1 - w)/w$, 所以, 当 $z = z_{cr}$ 时, $P_x = P_y$. 许多参与者只参与一回合博弈. 因此, 表达式 (4.37) 不成立. 但对于 w_n 的许多其他选择非常有效, 譬如假设参与者都经历 N 个回合.

4.6 正义的拒绝

到目前为止, 我们所讨论的条件策略看似矛盾, 且几乎不一致. 实际上, 为什么一个回报者会拒绝捐赠呢? 这会影响回报者自己的声誉, 从而降低其从其他回报者获得帮助的可能性. 回报者自己成了被报复的对象. 他们尽职的维护社区的秩序, 使得那些潜在的背叛者放弃坏念头. 但这样做, 自己却付出了代价. 他们怎么会选择这样的策略呢?

有一个显而易见的解释. 将正义的与非正义的背叛予以区分. 这种区分可以做到. 譬如, 类似于 3.17 节中讨论直接回报时采用的方法那样, 进行好坏声誉的标识. 参与者最初都具有一个好的声誉, 且要一直帮助好声誉的参与者, 其好的声誉将一直保持. 这些参与者即便在背叛时依然可以保持好的声誉, 只要其所背叛的对手是具有坏声誉的参与者.

然而声誉的标识是相当复杂的, 它需要对整个群体进行持续监控, 这对参与者而言负担过重. 事实上, 假设你的接受者 A 在前一回合拒绝帮助接受者 B, 这个拒绝是正义的吗? 如果 B 被证明帮助了所有拥有好声誉的对手, 那么 A 的拒绝显然是非正义的. 但如果 B 曾拒绝过参与者 C, 那又该如何判断呢? 于是, 你还必须知道参与者 B 对 C 的背叛是否是正义的. 这意味着你需要了解参与者 C 的历史博弈等等. 对于直接回报而言, 把握声誉要容易得多, 只需要明了对手与自己在前面的回合博弈时的策略选择. 即使这样, 误判也会导致死锁: 有可能两个参与者都认为自己拥有好声誉, 并坚定地惩罚对方. 在间接回报中, 问题变得更为严峻, 参与者不仅需要跟踪当前接受者的历史策略, 还需要跟踪该对手在之前博弈中的对手的历史策略等.

对于个体而言, 应该如何通过预判来更新对手的形象、评价或声誉, 这一点并不明显. 事实上, 导致社会规则最终形成的道德评价也在不断进化.

在后续关于间接回报的章节中, 假设个体能够跟踪群体中所有成员声誉的变化 (因此, $q = 1$), 并且假设在充当捐赠角色时, 其决定帮助与否主要依赖于对手的声誉, 偶尔也考虑自己的声誉. 这可以视为一个对局部信息的简单处理机制的探讨. 但它对社会规范的演化有着广泛深远的影响, 并进而影响道德判断的进化. 这涉及如下的决定性问题: 背叛何时正义, 何时非正义. 换句话说, 参与者何时为好, 何时为坏?

首先, 我们在一个非常特定的背景下考虑这个问题. 假设声誉或评价取值只有两种可能: 好或坏. 当然, 也可以设计许多其他策略. 例如, 考虑接受者和捐赠者的累计收益, 社区内合作的普遍程度, 或者某个参与者在前一回合是否获得帮助等等.

4.7 二值模型: 黑白世界

我们将考虑基于评价模式和行为模式的策略. 参与者 C 一旦观察到参与者 A 与参与者 B 之间的博弈, 其评价模式将启动 (图 4.1 中所描述替代的回报). 在观察者 C 的眼中, 潜在的捐赠者 A 的声誉可能会受到其决策行为的影响, 而潜在的接受者 B 作为博弈过程中被动的一方, 其声誉保持不变. 当参与者扮演潜在的捐赠者时, *行为模式* (action module) 决定其是否提供帮助. 这个决定依赖于参与者的评价模式所提供的信息.

我们现在讨论评价模式, 为简化起见, 假定在个体 C 的眼中个体 A 的声誉, 只依赖于 C 最近一次观察到 A 在博弈中作为捐赠者时的策略行为: 即是否帮助第三方 B. 因此, C 的记忆非常有限, 并且对 A 的评价 (在 C 看来) 只能取好或坏两值之一. 对于 C 所观察到的每一次交互, 存在两种可能的结果 (A 提供或拒绝帮助), A 的声誉可能有两种, B 也如此. 于是, 共有 8 种可能的交互, 又依赖于 C 的认可

与否, 这样就有 $2^8 = 256$ 种不同的评价模式或价值系统.

举一个直观有趣的评价模式的例子, 我们考虑表格 4.1 中的三个价值系统——所谓的道德系统雏形, 分别记为标记, 声誉和判断. 对于所观察到的行为, 在不认可哪些行为并赋以 "坏" 值方面, 这些评价模式存在差异. 若参与者采用标记的评价模式, 则总是反对潜在的捐赠者拒绝帮助潜在的接受者. 4.2 节和 4.3 节中的回报者就是采用此评价模式, 它不考虑与被评价者进行交互的参与者的声誉. 相比之下, 采用声誉评价模式的参与者会谴责那些拒绝帮助*好*声誉参与者的行为, 但会宽容那些拒绝帮助*坏*声誉参与者的行为. 而采用判断评价规则的参与者还会进一步将帮助*坏*声誉参与者的行为纳入谴责的行列, 并且认为*坏*声誉参与者无法通过拒绝帮助其他*坏*声誉参与者来而获得好的声誉.

因此, 这三类价值系统对待坏行为的宽严程度不同. 简而言之, 在标记型评价者的眼中, 任何拒绝帮助对手的参与者都是坏的. 而在声誉型评价者的眼中, 参与者只有在拒绝帮助*好*声誉的对手时才会被认为是坏的. 在判断型评价者的眼中, 拒绝帮助*好*声誉的对手以及帮助*坏*声誉的对手都会被认为是坏的 (表 4.1).

表 4.1 评价模式

策略/情形	标记 (scoring)	声誉 (standing)	判断 (judging)
好 → 好	好	好	好
好 → 坏	好	好	坏
坏 → 好	好	好	好
坏 → 坏	好	好	坏
好 ↛ 好	坏	坏	坏
好 ↛ 坏	坏	好	好
坏 ↛ 好	坏	坏	坏
坏 ↛ 坏	坏	好	坏

注: 该模式表明被观测的博弈中潜在的捐赠者的声誉值 (好 → 坏表示 "好声誉的参与者帮助坏声誉的参与者." 坏 ↛ 好表示 "坏声誉的参与者拒绝帮助好声誉的参与者", 等等)

可以对评价模式分类: 如果评价只依赖于所观察到的行为, 则称之为*一阶的*; 进一步, 如果还依赖于接受者的声誉, 则称之为*二阶的*; 再进一步, 如果还依赖于捐赠者的声誉, 则称之为三阶的. 标记为一阶的, 声誉为二阶的, 而判断为三阶的.

现在考虑行为模式, 假设捐赠者决定是否提供帮助完全取决于交互中的捐赠者及接受者双方的声誉. 因为存在四种情形 (捐赠者和接受者可以分别为好或坏), 所以共有 $2^4 = 16$ 种决策规则. 四类直观有趣的例子为 Co, Self, And 和 Or (表 4.2). Co 类型的参与者的决策只受潜在接受者的声誉影响, 当且仅当接受者为好声誉时才提供帮助. 目前为止所考虑模型中的回报者就是采取这种行为模式. Self 类型的参与者只关心自己的声誉, 当且仅当自己为坏声誉时才会提供帮助. And 类型的参

与者只有在接受者为好声誉且自己为坏声誉时才提供帮助, 而 Or 类型的参与者在接受者为好声誉或者自己为坏声誉时会提供帮助. 当然, 16 种决策规则中也包含两类无条件规则, 即总是提供帮助和总是拒绝帮助的 AllC 和 AllD, 这两种类型的决策完全不依赖于声誉.

表 4.2　行为模式

情形/策略	Self	Co	And	Or	AllC	AllD
好 $\xrightarrow{?}$ 好	否	是	否	是	是	否
好 $\xrightarrow{?}$ 坏	否	否	否	否	是	否
坏 $\xrightarrow{?}$ 好	是	是	是	是	是	否
坏 $\xrightarrow{?}$ 坏	是	否	否	是	是	否

注: 该模式用于根据参与者自己和潜在接受者的声誉来决定是否提供帮助 (若参与者采用此行为模式, 坏 $\xrightarrow{?}$ 好将决定其在拥有坏声誉的情形下遇到好声誉的接受者时是否提供帮助, 等等)

在间接回报模型中, 每一个策略由具体的评价模式和行为模式组合而成. 因此, 总共有 $2^4 \times 2^8 = 2^{12} = 4096$ 种策略. 它们之间并非各不相同, 譬如所有行为采用 AllD 模式的策略, 不管评价模式如何, 它们的效果都是相同的.

4.8　前　八　名

然而策略还是显得过多. 为简化起见, 假设每个参与者的声誉是众所周知的, 这意味着群体中采用统一的评价模式, 并且 $q = 1$: 每位参与者都知道对手的声誉. 此时, 仅有 8 类回报策略是合理的, 此处的合理性基于以下意义.

(a) 对采用单一策略的群体, 只要不等式 $wb > c$ 成立, 则少量采用不同行为模式 (但采用相同的评价模式) 的参与者不能侵入.

(b) 在这类群体中, 每一回合的平均适合度接近理论的最大值 $b - c$ (或者最多相差一个 ϵ 阶的项, ϵ 为捐赠意图因误操作而未成功实施的概率).

前八名的这些策略由表 4.3 列出.

只有 Co 以及 Or 类型的行为模式在前八名中出现. 这类参与者总会帮助好声誉的参与者, 并且当自己为好声誉且对方为坏声誉时, 总是拒绝帮助. 前八名策略的评价模式存在如下的一致性: 无论参与者的声誉如何, 他们只根据该参与者是否帮助好声誉的接受者来判定其好坏与否. 此外, 它们都认可好声誉的参与者背叛坏声誉的参与者.

对于以下三类行为: 好声誉的参与者帮助坏声誉的参与者, 坏声誉的参与者帮助及背叛坏声誉参与者, 前八名策略对它们的评判是待定的. 这 2^3 种不同的取值恰好构成了前八名策略. 如果评价模式要求坏声誉的参与者帮助坏声誉的参与

者, 则对应的行为模式为 Or, 而其余为 Co. 注意到, 前八名中包含评价模式为声誉和判断的策略, 而没有标记的策略. 特别地, 我们在 4.1 节中讨论的简单的策略 "Co-Scoring", 并未出现在前八名中.

表 4.3 前八名策略

情形/策略	L1	L2	L3	L4	L5	L6	L7	L8
好 → 好	好	好	好	好	好	好	好	好
好 → 坏	好	坏	好	好	坏	坏	好	坏
坏 → 好	好	好	好	好	好	好	好	好
坏 → 坏	好	好	好	坏	好	坏	坏	坏
好 ↛ 好	坏	坏	坏	坏	坏	坏	坏	坏
好 ↛ 坏	好	好	好	好	好	好	好	好
坏 ↛ 好	坏	坏	坏	坏	坏	坏	坏	坏
坏 ↛ 坏	坏	坏	好	好	好	好	坏	坏
好 $\xrightarrow{?}$ 好	是	是	是	是	是	是	是	是
好 $\xrightarrow{?}$ 坏	否	否	否	否	否	否	否	否
坏 $\xrightarrow{?}$ 好	是	是	是	是	是	是	是	是
坏 $\xrightarrow{?}$ 坏	是	是	否	否	否	否	否	否

注: 评价模式 (表中前 8 行) 和行为模式 (最后 4 行) 定义了每个策略. 这些策略拥有最高的收益, 并且可以成功抵抗住背叛者的侵入. 策略 L3 对应于 Co 声誉 (Co-Standing), 而策略 L8 对应于 Co 判断 (Co-Judging). 表中未出现含有标记评价模式的策略. 待定的评价包含三类情形: 好 → 坏, 坏 → 坏和坏 $\xrightarrow{?}$ 坏. 这八类评价模式分别对应于不同的行为模式, 行为模式可以为 Or 或者 Co

4.9 解释前八名

首先我们不考虑出错. 条件 (a) 和 (b) 意味着在同质群体中, 参与者总是合作, 而 AllC 和 AllD 参与者都不能侵入. 实际上, 可以得出采用其他行为准则的参与者也不能侵入.

显然, 参与者的成功与否完全由其对原住参与者所采取的行为决定. AllD 参与者由于背叛原住参与者, 虽然节约了成本 c, 但也丧失了获得较高的总收益的机会 (否则, AllD 参与者可以侵入). 这意味着根据条件 (a) 原住参与者总是相互帮助, 从而必定得到补偿. 这只有在原住参与者下一回合获得帮助的可能性较高, 即背叛者不会获得帮助 (收益为 0) 时才发生. 这是由于下一回合发生的概率为 w, 原住参与者的期望收益为 wb. 因此, 为了使原住参与者获得补偿, 条件 $c < wb$ 必须满足.

只有在可以获得新声誉的情况下, 才可以将背叛者与帮助其他原住参与者的原

住参与者加以区分. 通常背叛者被标记为坏人, 而帮助其他原住参与者的回报者被标记为好人. 这蕴含着两类行为规则: 好 $\stackrel{?}{\to}$ 好: 是, 好 $\stackrel{?}{\to}$ 坏: 否 (表 4.3 中第 9 行和第 10 行). 对于一个提供帮助的行为规则而言, 我们称其是有利的: 只有在遵循它时才可获得好声誉, 反之则获得坏声誉. 因此, 表 4.3 中第 9 行蕴含了第 1 行和第 5 行. 类似的, 一个拒绝帮助的行为规则是有利的, 若违反该规则不会获得更好的声誉值. 特别地, 第 10 行意味着或者 (i) 好 $\nrightarrow$ 坏是好的 (行 6), 或者 (ii) 好 $\nrightarrow$ 坏与, 好 $\to$ 坏都是坏的. 但第 (ii) 项意味着坏的是可以传染的: 好的参与者只要遇到坏的参与者, 无论采取何种行为都会得到坏声誉. 可见我们应该放弃 (ii).

现在考虑实施捐赠的过程中可能出现失误. 这类失误发生后, 该参与者被标记为坏, 且在下一回合不会获得原住参与者的帮助. 如果坏具有传染性, 则坏的参与者的数量不会减少. 更糟糕的是, 任何一个回合都可能失误, 所以坏的参与者的数量会增加.

如果无法摆脱坏的声誉, 那么因误操作而背叛的原住参与者将不再有任何收益. 如果还有很多回合, 这意味着大多数参与者最终将不再获得收益. 因此必须得有一个行为使坏的捐赠者能够摆脱其现有的声誉. 因为与其他类型相遇的概率比较小, 所以此行为必定是针对好的原住参与者. 且该行为只能是坏 $\stackrel{?}{\to}$ 好: 是, 否则与 AllD 参与者毫无差别. 所以, 这反过来蕴含了坏 $\to$好被视作好, 而坏$\nrightarrow$好则被视为坏 , 即第 11 行蕴含着第 3 行和第 7 行. 我们已经确定了行为模式中的三行 (第 9, 10 及 11 行) 以及评价模式中的五行 (第 1, 3, 5, 6 及 7 行). 评价模式中其余三行, 即第 2, 4 及 8 行待定. 有 2^3 种方法去代入这三行. 容易知道, 对于所得的八类评价模式中的任意一种, 都恰好存在一类行为模式, 与之组合形成的策略满足前一节中的条件 (a) 和 (b).

例如, 考虑 L1 的评价模式, 因为坏$\to$ 坏为好, 而坏 $\nrightarrow$ 坏 为坏, 所以明显的, 最优行为规则要求坏参与者帮助坏的参与者, 这就得到 Or 行为规则. L2 的评价模式也是如此. 因为好参与者总是不帮助坏参与者, 其相应的评价规则是无关紧要的. 相反的, 仅对表 4.3 中的前 2 列评价模式, Or 才是优于 Co 的. 坏 $\nrightarrow$ 坏: 是 要求提供帮助, 即通过帮助可以得到好声誉, 而拒绝则依然保持坏的声誉. 对于其他列而言, 这是不满足的.

为了更好的理解前八名策略的共性, 下面给出它们都满足的性质.

(A) 维持合作. 这意味着好参与者帮助好参与者, 且被视为好 (第 9 行和第 1 行).

(B) 标识背叛. 好或坏参与者如果拒绝帮助好参与者, 将被视为坏 (第 5 行和第 7 行).

(C) 正义惩罚. 好参与者遇到坏参与者时, 不会因为拒绝帮助而被贴上坏的标签 (第 10 行和第 6 行). 这就排除了一阶评价模式, 比如标记.

(D) 接受道歉. 因失误而导致的背叛不应造成永恒的灾难. 因此, 坏参与者可以通过与好的参与者合作而重获好声誉 (第 11 行和第 3 行).

显然在由相互合作的回报者构成的原住参与者群体中, 评价规则必须与行为规则相对应. Co-标记不满足此条件. 如我们所见, 尽管会失去自己好的声誉, 好参与者仍然要克制住自己, 避免帮助坏参与者. 然而有趣的是这个要求在前八名策略中存在两个例外, 即表 4.3 中策略 L7 和 L8. 坏参与者相遇时无法改变坏声誉. 然而在采取此类策略的同质群体中, 坏参与者之间相遇的概率非常小.

在采用前八名策略之一的参与者和 AllC 参与者构成的群体中, 所有参与者都选择合作. 两类参与者都同样可能因失误而背叛, 从而失去好声誉. 但是当遇到坏声誉的参与者时, 回报者将选择背叛, 从而节约成本 c, 并且不会丧失自己的好声誉. 因此, 回报者的收益正比于 ϵ 增长.

4.10 二阶评价

在二阶评价中, 行为的判定只依赖于参与者是否提供帮助以及接受者声誉的好坏, 而不考虑捐赠者声誉的好坏. 这样可以得到 16 种不同的评价模式. 有两种位于前八名策略之中, 即 L3 和 L6 (表 4.3). L3 策略采取声誉评价模式 (除背叛好声誉的接受者之外, 任何行为都是好的). 而根据策略 L6 的评价模式, 背叛好声誉的接收者或帮助坏声誉的接收者都是坏的, 其余的行为都是好的. 这两类策略都采用的是 Co 行为模式.

首先, 采用与 4.2 节和 4.3 节相同的分析方法, 我们将策略 L6 与策略 AllC 及策略 AllD(它们的频度分别为 z, x, y) 放在一起考虑. 记 h 为参与者拥有好声誉的概率. 其组成为: (a) 遇到好的参与者且没有失误的, 或者是遇到坏的参与者且恰好失误 (概率为 $h(1-2\epsilon+\epsilon)$) 的 AllC 参与者; (b) 遇到坏的参与者 (概率为 $1-h$) 的 AllD 参与者; (c) 在前一回合遇到好的参与者且没有犯错的, 或者是遇到坏的参与者的 (概率为 $1-h\epsilon$) 的回报者. 因此,

$$h = \frac{1-\bar{\epsilon}x}{2y+z+\epsilon(2x+z)}. \tag{4.38}$$

当 $n \geqslant 1$ 时, 有

$$P_x(n) = -c\epsilon + b\epsilon[x + z(h(1-2\epsilon)+\epsilon)], \tag{4.39}$$

$$P_y(n) = b\epsilon[x + z(1-h)], \tag{4.40}$$

以及

$$P_z(n) = -c\epsilon h + b\epsilon[x + z(1-h)]. \tag{4.41}$$

对于初始回合, 也有类似结果. 由此, 去掉因子 $\epsilon(1-w)$ 可得

$$P_x - P_y = -c + wb\bar{\epsilon}z(2h-1), \tag{4.42}$$

$$P_x - P_y = -ch + wb\bar{\epsilon}hz \tag{4.43}$$

和

$$P_z - P_x = (1-h)(c + wb\bar{\epsilon}z) > 0. \tag{4.44}$$

由于最后一个表达式在 $\mathrm{int}S_3$ 内为正, 所有轨线都收敛到 $x=0$. 在此面上, 我们得到以 $\hat{z} = c/w\bar{\epsilon}b$ 为不稳定平衡点的双稳态, 如图 4.10 所示.

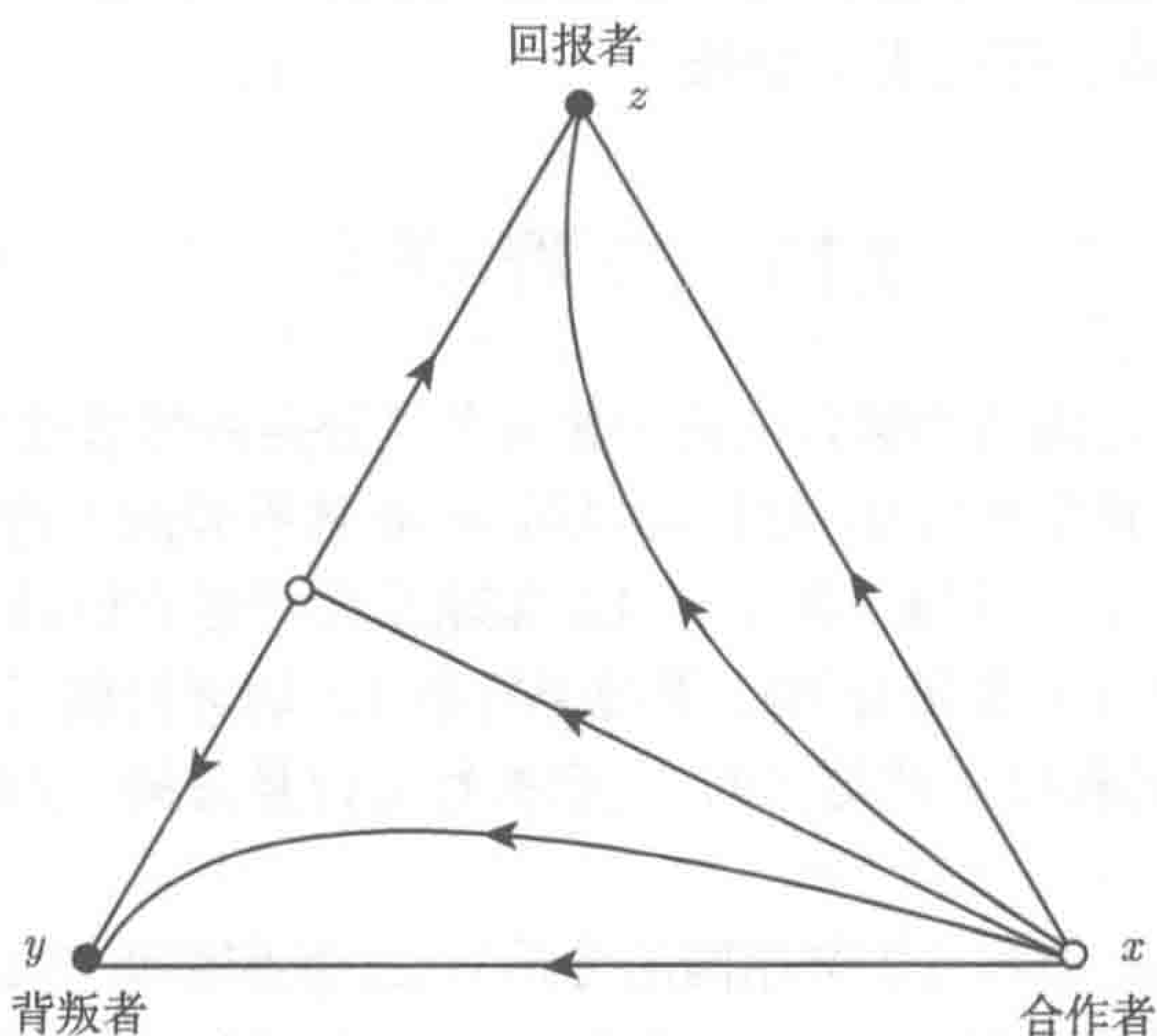


图 4.10　前八名策略中二阶评价模式的复制动力学行为

对于二阶策略 L3, 结论基本相同. 对于 AllC 参与者, 如果遇到好参与者且没有出现失误或者遇到坏参与者, 则会有好声誉: 其概率为 $1-h\epsilon$. 对于回报者, 结论同样成立. AllD 参与者只有在前一回合遇到坏参与者后才会被视为好, 概率为 $1-h$. 这样, 有

$$h = \frac{1}{1+y+\epsilon(x+z)}. \tag{4.45}$$

显然,

$$P_x(n) = -c\epsilon + b\epsilon[x + z(1-h\epsilon)], \tag{4.46}$$

$$P_y(n) = b\epsilon[x + z(1-h)], \tag{4.47}$$

以及

$$P_z(n) = -c\epsilon h + b\epsilon[x + z(1-h\epsilon)]. \tag{4.48}$$

同样, $P_z - P_x = c\bar{\epsilon}(1-h) > 0$ 意味着在单形 S_3 内 $x \to 0$, 并且在边界 $x = 0$ 上存在不稳定平衡点 $z = c/bw\bar{\epsilon}$.

如果考虑感知失误 (这表示大多数参与者知道正确的声誉, 存在 ϵ 的少数参与者会弄错), 那么原本是纯回报者的吸引子将变成由回报者与无条件合作者组成的混合吸引子. 这种假设适用于所谓的 "间接观察模型", 它假定 A 和 B 之间的互动只被一个观察者看到, 譬如 C, 而所有其他参与者都采纳 C 对 A 的评价. "直接观察模型" 的情形则更为复杂. 在这种模型中, 所有参与者各自确定并记录对手声誉值. 本质上, 评价模式应该在这个框架下讨论. 有人认为, 正是由于语言, 才使得群体的所有成员在声誉的评价上保持一致, 闲聊是非常有效的途径, 它使得每个人都能获得所有历史回合的信息. 但我们都有同感, 即便两个参与者都直接观察了同一博弈, 他们也可能得出截然不同的评价. 所以应该考虑私有声誉值.

4.11 参考文献

间接回报出自 Sugden (1986) 及 Alexander (1987). 逆向回报及顺向回报的划分则出自 Boyd 和 Richerson (1989). 在经典的博弈理论中, 与不同对手进行重复博弈的问题已经由 Rosenthal (1979), Kandori (1992) 及 Okuno-Fujiawara 和 Postlewaite (1995) 等人作过研究. 借助进化博弈理论研究间接回报的方法则由 Nowak 和 Sigmund (1998a,1998b) 提出, 也可参见 Ferriere (1998). 失误在稳定合作中所起的重要作用则由 Lotern, Fishman 和 Stone (1999) 指出. Ohtsuki (2004) 研究了间接回报的反应策略. 4.2 节中的持续进入模型由 Brandt 和 Sigmund (2005) 提出. Fishman (2003), Mohtashemi 和 Mui (2003), 以及 Brandt 和 Sigmund (2004, 2005) 等文章研究了将 AllC 和回报者策略之间的平衡点变为吸引子的各种方法. 允许正义背叛的各种评价规则由 Sugden (1986), Nowak 和 Sigmund (1998a), Leimar 和 Hammerstein (2001), Panchanathan 和 Boyd (2003), 以及 Takahashi 和 Mashima (2003[①]) 等提出, 也可参见 Nowak 和 Sigmund (2005). 由于术语并不统一, "Standing" 在不同作者的文献中含义不同. 评价模式和行为模式的分类由 Brandt 和 Sigmund (2004) 以及 Ohtsuki 和 Iwasa (2004) 提出. 前八名策略最初由 Ohtsuki 和 Iwasa (2004) 引入, 进一步的研究见 Ohtsuki 和 Iwasa (2006). Ohtsuki 和 Iwasa (2007) 研究了二阶评价的动力学行为. 更多关于错位回报的讨论见 Bshary 和 Grutter (2006), Engelmann 和 Fischbacher (2002), Pfeiffer 等 (2005), Nowak 和 Roch (2007), Rutte 和 Taborsky (2007). Suzuki 和 Akiyama (2007a, 2007b) 以及 Masuda 和 Ohtsuki (2007) 研究了大规模群体内的间接回报. 更多关于二值规范的讨论, 可参见 Pacheco, Santos 和 Chalub (2006) 以及 Chalub, Santos 和 Pacheco (2006). Bolton, Katok, 和 Ocken-

①译注: 原文此处为 "2004".

fels (2004, 2005[①]) 以及 Keser (2003) 通过实验研究了声誉机制的有效性; Lotem, Fishman 和 Stone (2003[②]) 研究了通过释放利好消息的无条件利他主义者的演化. Milinski 等 (2001) 作了对于不同评价模式的实验性研究. Dufwenberg 等 (2001) 以及 Roberts (2008) 对直接和间接回报作了比较, 也可参见 Pollock 和 Dugatkin (1992).

①译注: 原文此处为 “2004a, 2004b”.
②译注: 原文此处为 “2002”.

第 5 章　公平和信任: 激励的力量

5.1　最终的报价

在最后通牒博弈中 (1.11 节), 两个参与者被随机指定为提议者或响应者. 然后实验者提供一定的金额给提议者. 提议者再把其中一部分金额分给响应者. 如果响应者接受, 则金额分配随之确定, 博弈结束. 如果响应者拒绝, 则全部金额返还实验者, 博弈也结束: 但此时, 两个参与者什么都得不到. 需要强调的是参与者双方预先知道这个规则, 并且知道他们不会再次相遇.

我们将金额规范化为 1, 且用 p 表示报价的大小. 那么, 提议者的策略可以简单规定为 $p \in [0,1]$. 响应者的策略可以用其可接受的报价集合来表示, 比如区间 $[q,1]$. 因此, 响应者的策略是由他的期望水平 $q \in [0,1]$ 给出的. 如果 $p \geqslant q$, 那么响应者接受报价且获得收益 p, 提议者获得 $1-p$. 如果 $p < q$, 则报价被拒绝, 双方收益都为零.

在实验中, 大多数的报价在 40% 到 50% 之间, 且都能被接受. 少数的报价低于 20%, 通常被拒绝. 大部分提议者似乎预料到这种情况, 这促使他们给出一个合理的报价. 但一个利己的响应者会接受任意正的报价, 因为这总比什么都没有强. 那么, 为什么期望水平通常都高于 20% 呢? 提议者与响应者是两个不同的角色, 我们分别用 I 和 II 表示. 参与者的策略集合 (理论上) 在单位区间 $[0,1]$ 上是连续的. 当然, 实际中, 由于货币面额的离散特性, 集合是有限的. 甚至我们可以简单的假设, 提议者只能在报价 h 和 l 之间选择 $(0 < l < h < 1)$. 例如, 高报价 h 为全部的 40%, 低报价 l 为全部的 15%. 原则上, 响应者可以接受两种报价, 或其中一种, 或都不接受. 我们再假设响应者只在两种策略之间做选择: 一种策略只接受高报价, 另一种策略接受两种报价.

因此, 我们得到一个简化的模型. 对角色 I, 策略 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 分别由报价 h 和 l 表示; 对角色 II, 为了方便起见, 策略 $\boldsymbol{f}_1$ 和 $\boldsymbol{f}_2$ 仍分别由 h 和 l 表示: 此为响应者的期望水平. 收益矩阵为

$$
\begin{array}{c|cc}
 & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\
\hline
\boldsymbol{e}_1 & (1-h,h) & (1-h,h) \\
\boldsymbol{e}_2 & (0,0) & (1-l,l)
\end{array} . \tag{5.1}
$$

这个两角色博弈有两个明显的纳什均衡 (h,h) 和 (l,l).

在分析相应的群体动力学之前, 我们注意到, 实际上最后通牒是一个对称博弈. 这在抛硬币决定谁是提议者, 谁是响应者之后成立. 因此, 最后通牒的策略为 $(p, q) \in [0,1]^2$. 而对于简化的最后通牒博弈, 策略为 $(\boldsymbol{e}_i, \boldsymbol{f}_j)$.

5.2 迷你博弈的迷你课程

在考虑简化的对称最后通牒博弈之前, 有必要先分析更一般的两角色博弈的群体动力学. 考虑有两角色 I 和 II 的博弈, 且每个角色有两个策略, 用 $\boldsymbol{e}_i$ 表示角色 I 的策略, $\boldsymbol{f}_j$ 表示角色 II 的策略 $(i, j \in \{1, 2\})$. 收益矩阵可以表示为

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (A, a) & (B, b) \\ \boldsymbol{e}_2 & (C, c) & (D, d) \end{array}. \tag{5.2}$$

如同 2.5 节, 我们考虑的对称模型 (抛硬币决定参与者的角色). 对于所得的对称博弈, 其策略表示为 $\boldsymbol{G}_1 = \boldsymbol{e}_1\boldsymbol{f}_1, \boldsymbol{G}_2 = \boldsymbol{e}_2\boldsymbol{f}_1, \boldsymbol{G}_3 = \boldsymbol{e}_2\boldsymbol{f}_2$ 和 $\boldsymbol{G}_4 = \boldsymbol{e}_1\boldsymbol{f}_2$. 与采用策略 $\boldsymbol{G}_j$ 的参与者对局时, 采用策略 $\boldsymbol{G}_i$ 的参与者收益由下面矩阵中的 (i, j) 元素给出 (其中的因子 1/2 被省略)

$$\boldsymbol{M} = \begin{pmatrix} A+a & A+c & B+c & B+a \\ C+a & C+c & D+c & D+a \\ C+b & C+d & D+d & D+b \\ A+b & A+d & B+d & B+b \end{pmatrix}. \tag{5.3}$$

这对应方程 (2.14). 例如, $\boldsymbol{G}_1$ 参与者与 $\boldsymbol{G}_3$ 参与者对局时, 有 1/2 的概率扮演角色 I, 此时采用策略 $\boldsymbol{e}_1$ 应对采用 $\boldsymbol{f}_2$ 策略的对手, 获得收益 B; 有 1/2 概率扮演角色 II, 此时采用策略 $\boldsymbol{f}_1$ 应对采用策略 $\boldsymbol{e}_2$ 的对手, 获得收益 c.

复制系统

$$\dot{x}_i = x_i[(\boldsymbol{Mx})_i - \boldsymbol{x} \cdot \boldsymbol{Mx}] \tag{5.4}$$

描述了状态 $\boldsymbol{x} = (x_1, x_2, x_3, x_4) \in S_4$ 的进化. 因为将 m_{ij} 用 $m_{ij} - m_{1j}$ 替换 $(i, j \in \{1, 2, 3, 4\})$, 系统不会受影响, 所以我们采用矩阵

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ R & R & S & S \\ R+r & R+s & S+s & S+r \\ r & s & s & r \end{pmatrix}, \tag{5.5}$$

其中 $R := C - A$, $r := b - a$, $S := D - B$ 及 $s := d - c$. 仍将此矩阵记为 $\boldsymbol{M}$, 其满足

$$m_{1j} + m_{3j} = m_{2j} + m_{4j}, \quad j = 1, 2, 3, 4. \tag{5.6}$$

因此

$$(\boldsymbol{Mx})_1 + (\boldsymbol{Mx})_3 = (\boldsymbol{Mx})_2 + (\boldsymbol{Mx})_4 \tag{5.7}$$

对所有的 $\boldsymbol{x}$ 都成立. 由此性质和方程 (2.32) 可以得出函数 $V = x_1x_3/x_2x_4$ 在 S_4 的内部满足

$$\dot{V} = V[(\boldsymbol{Mx})_1 + (\boldsymbol{Mx})_3 - (\boldsymbol{Mx})_2 - (\boldsymbol{Mx})_4] = 0. \tag{5.8}$$

所以, V 是复制动力系统的一个不变运动: 沿每条轨道, 其值不变.

这样, 单形 S_4 的内部是由如下层状的面构成

$$W_K := \{\boldsymbol{x} \in S_4 : x_1x_3 = Kx_2x_4\}, \tag{5.9}$$

其中 $0 < K < \infty$. 这些鞍面是由四边形的边 $\boldsymbol{G}_1\boldsymbol{G}_2$, $\boldsymbol{G}_2\boldsymbol{G}_3$, $\boldsymbol{G}_3\boldsymbol{G}_4$, $\boldsymbol{G}_4\boldsymbol{G}_1$ 和单形 S_4 的顶点张成, 如图 5.1 所示.

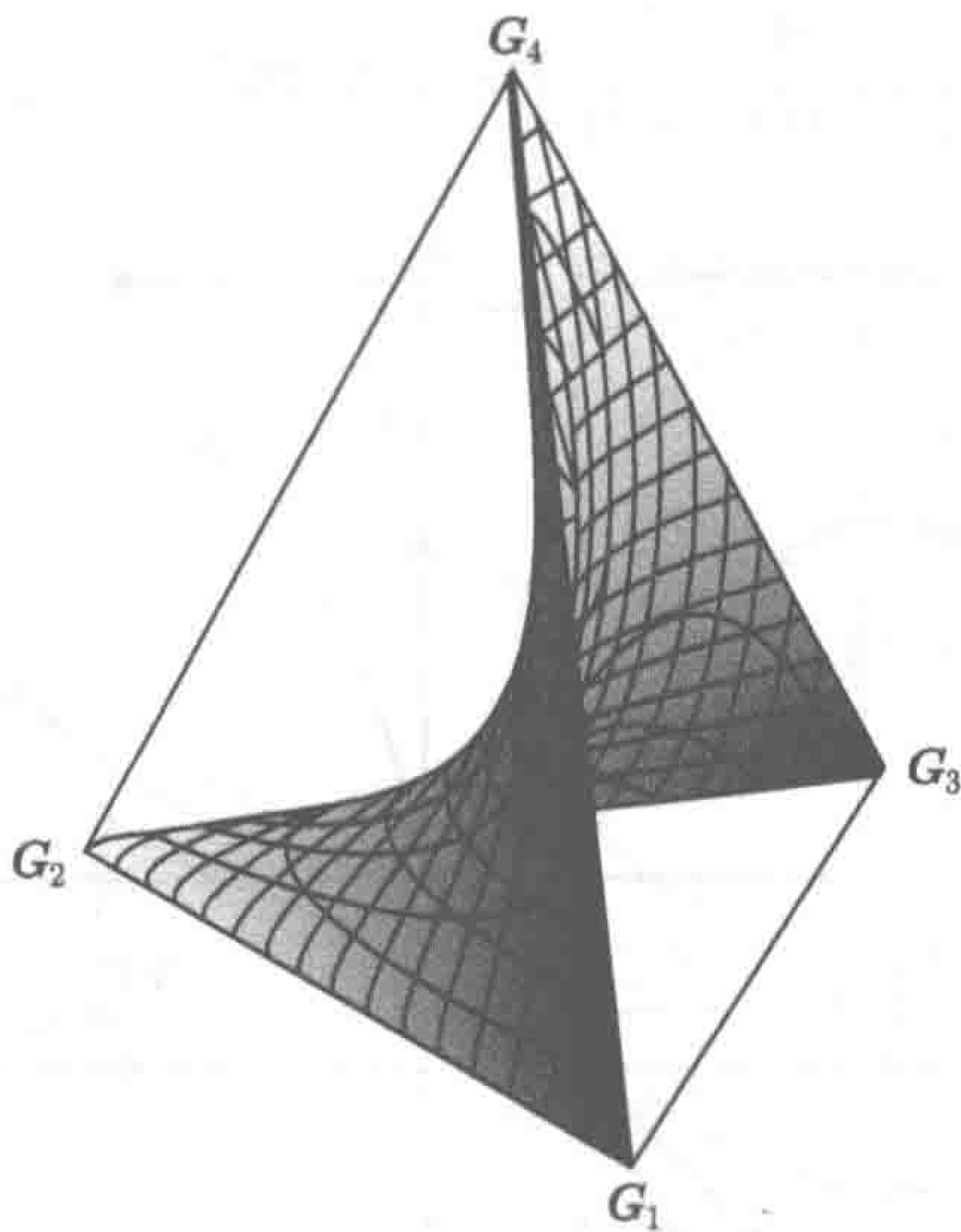


图 5.1 由边 $\boldsymbol{G}_1 - \boldsymbol{G}_2 - \boldsymbol{G}_3 - \boldsymbol{G}_4 - \boldsymbol{G}_1$ 所张成的单形 S_4 (四面体), 它由一层层的鞍状面 $W_K = \{\boldsymbol{x} \in S : x_1x_3 = Kx_2x_4\}$ 组成. 对两角色 (每个角色具有双策略) 博弈对称化后的复制系统, 这些面都是不变的

流在边界上的走向很容易由前面的矩阵得到. 例如, 当 $R = 0$ 时, 边界 $\boldsymbol{G}_1\boldsymbol{G}_2$ 由不动点组成. 当 $R > 0$ 时, 流沿着边界从 $\boldsymbol{G}_1$ 到 $\boldsymbol{G}_2$ (这意味着无策略 $\boldsymbol{G}_3$ 和 $\boldsymbol{G}_4$ 时, 策略 $\boldsymbol{G}_2$ 对 $\boldsymbol{G}_1$ 占优), 反过来, 当 $R < 0$ 时, 流沿着边界从 $\boldsymbol{G}_2$ 到 $\boldsymbol{G}_1$.

一般的, 参数 R, S, r 和 s 都不为零. 这对应四边形 $\boldsymbol{G}_1\boldsymbol{G}_2\boldsymbol{G}_3\boldsymbol{G}_4$ 的 16 种相图, 利用等价性可以约化为 4 个, 如图 5.2 所示. 因为 $(\boldsymbol{Mx})_1$ 为零, 单形 S_4 的内部驻

点必须满足 $(\boldsymbol{Mx})_i = 0, i = 2, 3, 4$. 这意味着对 $S \neq R$, 有

$$x_1 + x_2 = \frac{S}{S - R}, \tag{5.10}$$

而对 $s \neq r$, 有

$$x_1 + x_4 = \frac{s}{s - r}. \tag{5.11}$$

这些解位于单形上的充分必要条件是 $RS < 0$ 且 $rs < 0$, 对应张成 W_K 面的四边形 (c) 和 (d) 的相图. 这样, 我们得到由驻点组成的一条线, 且它与每个 W_K 面正好交于一点 (图 5.1). 这些解可以写成

$$x_i = m_i + \xi, \tag{5.12}$$

其中 $i = 1, 3$, 且

$$x_i = m_i - \xi, \tag{5.13}$$

其中 $i = 2, 4$, ξ 为参数, 且

$$\boldsymbol{m} = \frac{1}{(S - R)(s - r)}(Ss, -Sr, Rr, -Rs) \in W_1. \tag{5.14}$$

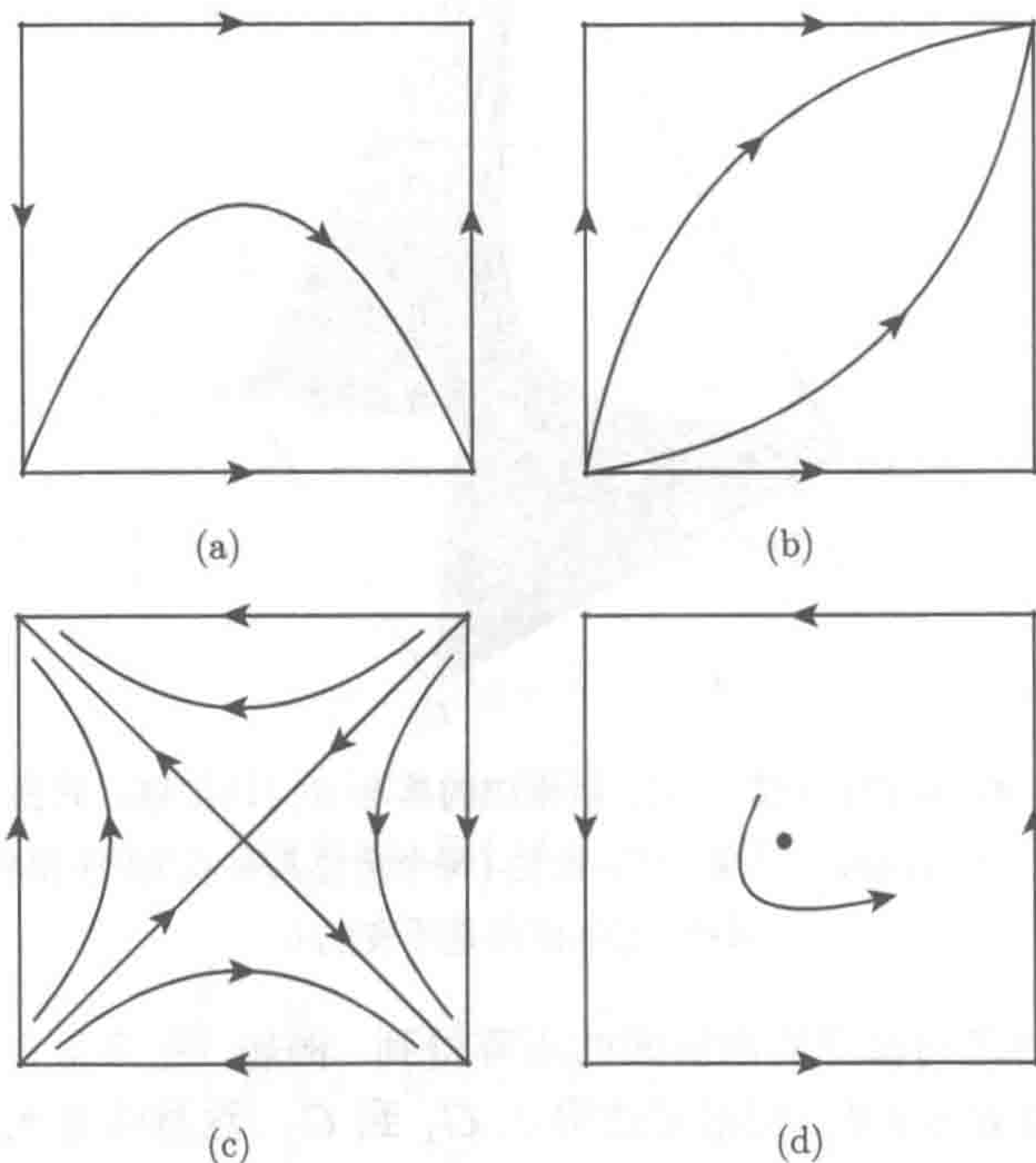


图 5.2 边界 $G_1 - G_2 - G_3 - G_4$ 上流的四种走向 (在等价意义下). 情形 (a) 和 (b), 正方形内部没有驻点, 因此面 W_K 也没有. 情形 (c) 和 (d), 存在驻点

5.3 回到简化的最后通牒

现在回到简化的最后通牒博弈. 策略 $\boldsymbol{G}_1$ 对应 (h,h): 高报价, 高期望水平. 可以认为这是*公平* (fair) 或*亲社会* (pro-social) 策略. 相比之下, $\boldsymbol{G}_3=(l,l)$ 代表了*利己* (asocial) 策略. 它接受任意正报价, 且提供的报价尽可能低. 策略 $\boldsymbol{G}_2=(l,h)$ 是*矛盾的* (paradoxical): 它提供的少, 但期望高报价, 这明显有悖于康德的绝对命令. 最后, $\boldsymbol{G}_4$ 提供高报价, 接受低报价. 因为缺乏更好的词, 我们将其称为*温和的* (mild) 策略. 收益参数满足 $A=B=1-h, C=c=0, D=1-l, a=b=h$ 和 $d=l$. 因此, $R=h-1<0, r=0, S=h-l>0$ 且 $s=l>0$. 利己策略相对温和策略与矛盾策略都占优; 矛盾策略对亲社会策略与利己策略来说是劣势策略; 在没有其他两个策略的时候, 温和策略与亲社会策略是等价的: 确实, 所有的报价都是公平的. S_4 的内部不存在驻点. 事实上, 无论是 $x_2>0$ 还是 $x_3>0$, 我们都有 $(\boldsymbol{Mx})_4>(\boldsymbol{Mx})_1$, 因此, 比率 x_4/x_1 和 x_3/x_2 都是增长的. 在每一个面 W_K 上, 解流如图 5.3 所示.

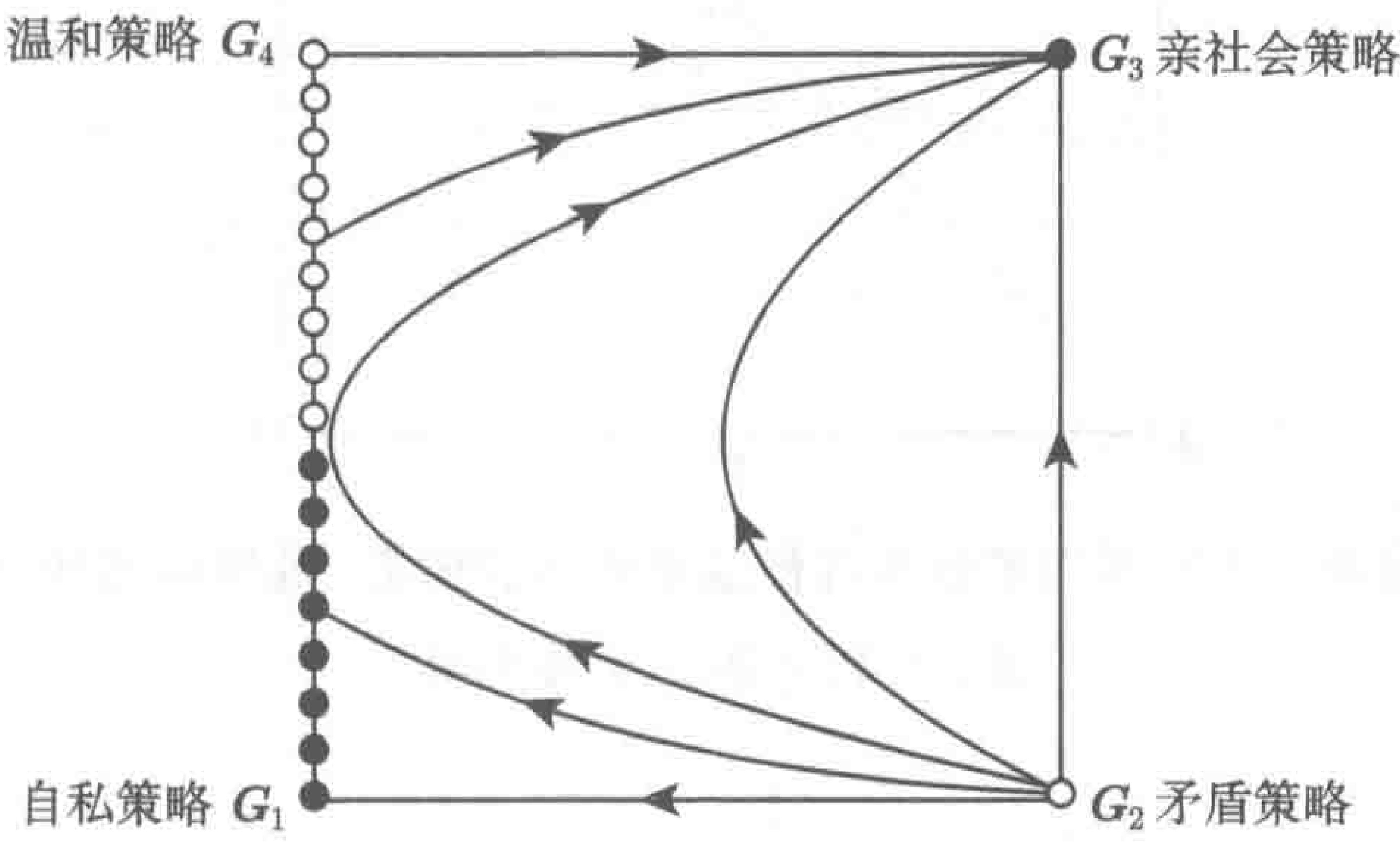


图 5.3 简化的最后通牒博弈 W_K 面上的复制动力学. 边 $\boldsymbol{G}_1\boldsymbol{G}_4$ 由驻点组成. 最终, 进化会趋向利己状态 $\boldsymbol{G}_3$

在边 $x_2=x_3=0$ 上, 所有的点都是驻点. 如果 $x_1<\dfrac{h-l}{1-l}$, 则 $(\boldsymbol{Mx})_2$ 和 $(\boldsymbol{Mx})_3$ 都大于 $\bar{M}=\boldsymbol{x}\cdot\boldsymbol{Mx}=0$. 点 $(\dfrac{h-l}{1-l},0,0,\dfrac{1-h}{1-l})$ 记为 $\boldsymbol{Q}$. 则博弈的对称纳什均衡为线段 $\boldsymbol{G}_1\boldsymbol{Q}$ 上的点和顶点 $\boldsymbol{G}_3$. 注意到, 边 $x_2=x_4=0$ 上还存在另一个驻点 P, 其坐标为 $(h,0,1-h,0)$. 在只有亲社会和利己参与者的群体中, 有一个双稳态竞争. 如果 $h<1/2$, 亲社会策略是风险占优的 (它具有更大的吸引盆).

S_4 内部的轨道或者趋于 $\boldsymbol{G}_3$, 或者趋于纳什均衡的线段. 如果假设随机冲击偶尔扰动群体状态, 则将在边 $x_2=x_3=0$ 上诱导中性漂变. 如果 $x_1<\dfrac{h-l}{1-l}$, 引入利

己策略 $\boldsymbol{G}_3$ 的随机扰动将导致 $\boldsymbol{G}_3$ 固定. 这意味着, 最终群体将只由利己的参与者构成. 因此, 进化博弈理论和经典博弈理论做出了相同的预测; 两者都与实验证据形成对照.

回到具有连续策略的完全最后通牒博弈是有趣的. 考虑基于个体的模拟, 从 1000 个个体的群体开始, 其策略 (p,q) 随机分布在策略正方形 $[0,1]^2$ 之中. 假设每个个体与随机选择的 50 个对手分别进行一个回合最后通牒博弈. 然后, 参与者可以通过模仿从群体中选择的榜样来更新策略, 榜样被模仿的概率与其总收益成正比. 进一步, 可以假设个体不时地从附近随机选择策略. 基于模仿和创新的模式会导致策略非常接近 $(0,0)$ (图 5.4), 即非常接近提供很少且接受任意报价的利己群体 – 这与人类普遍的公平准则相去甚远.

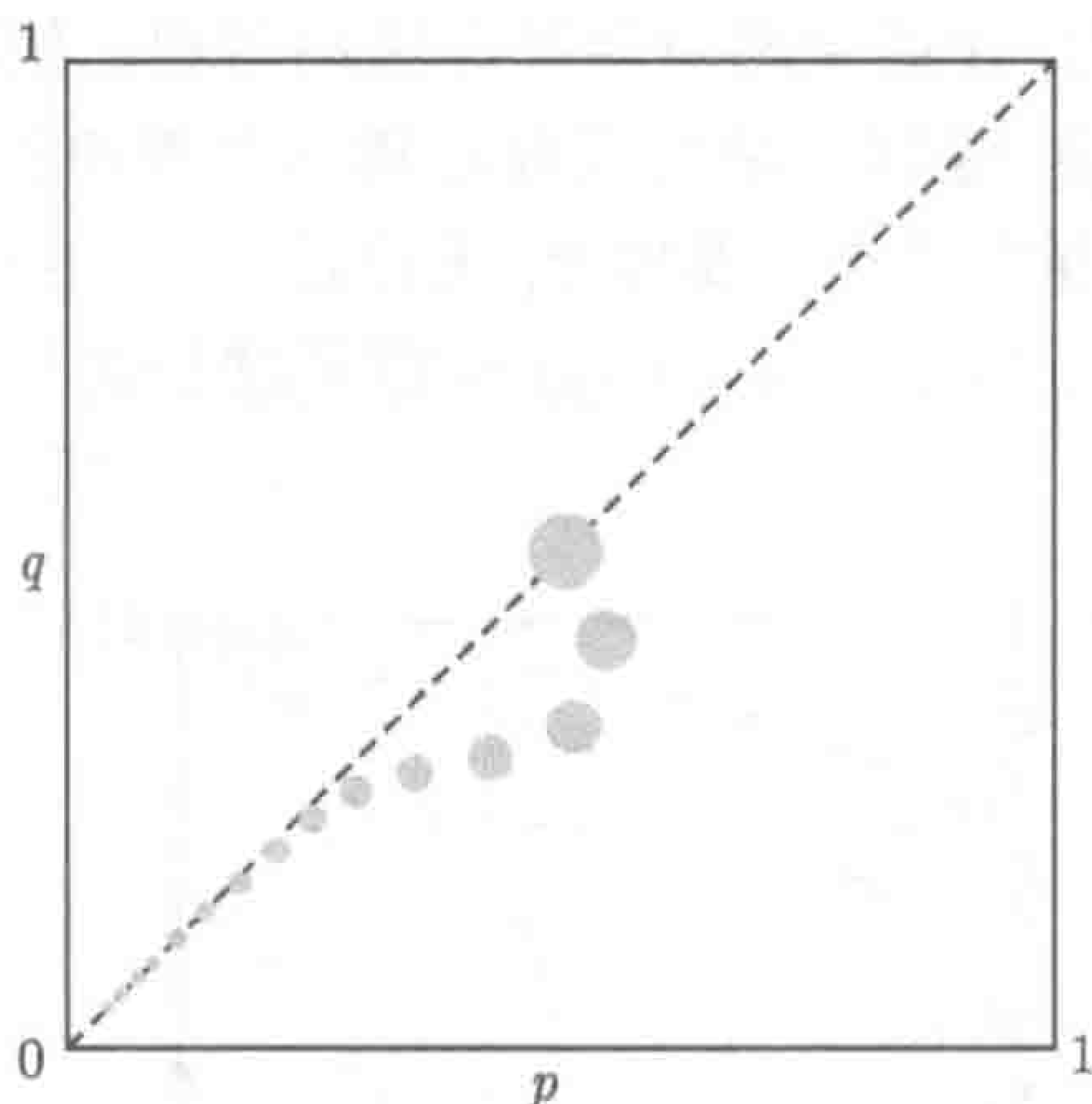


图 5.4 在严格匿名条件下对最后通牒博弈做基于个体的模拟, 其报价和期望水平的进化示意图. p 和 q 都趋于最小值

5.4 声誉产生分歧

迄今为止考虑的是严格匿名的条件. 现在假设有一定 (也许很小) 的概率, 参与者可能知道对手的声誉, 尤其是可能知道对手曾经接受过的报价. 进一步假设, 如果参与者有理由相信自己提供低报价能够得逞, 或更精确地, 知道对手以前接受过低报价, 那么参与者会比平时提供得更少. 这两个假设似乎足够合理: 只需了解群体中其他参与者的信息, 且具有少许投机取巧的利己主义者的动机. 在这种情况下, 参与者接受低报价将面临稍后遇到更低报价的风险.

为了分析这种情况, 我们再次回到只有两种报价 h 和 l 的简化最后通牒博弈. 假设 $\mu > 0$ 是公平的 (h,h) 提议者遇到温和的 (h,l) 响应者且知道响应者容易接受

低报价因此用报价 l 替换了 h 的概率. 这样就产生了如下的收益矩阵

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (1-h,h) & (1-h+\mu(h-l),h-\mu(h-l)) \\ \boldsymbol{e}_2 & (0,0) & (1-l,l) \end{array} . \tag{5.15}$$

它与矩阵 (5.1) 的差别只在 $(\boldsymbol{e}_1,\boldsymbol{f}_2)$ 位置. 这里, $\mu(h-l)$ 可以任意小, 其可视为先前博弈的一个扰动.

相应的对称博弈 (5.5) 其参数为 $R=h-1$, $r=-\mu(h-l)$, $S=(h-l)(1-\mu)$ 以及 $s=l$. 当 $\mu<1$ 时, 有 $R<0$, $S>0$, $s>0$ (如前) 以及 $r<0$①(在未受扰动的情形下 $r=0$). 这样, 对应于图 5.2 (c), 产生了通有情形. 状态空间 S_4 的内部存在一条由驻点构成的线. 每一个面 W_K $(K>0)$ 与这条线交于鞍点. 特别的, W_1 的驻点 (参考 3.6 节, 由于 $x_1x_3=x_2x_4$, 存在联接平衡) 由方程 (5.14) 给出,

$$\boldsymbol{m}=\frac{1}{k}(l(h-l)(1-\mu),(h-l)^2\mu(1-\mu),(h-l)(1-h)\mu,l(1-h)), \tag{5.16}$$

其中 $k=(1-l-\mu(h-l))(l+\mu(h-l))$. 当 $\mu\to0$ 时, $\boldsymbol{m}$ 和所有内部驻点, 趋于边界 $\boldsymbol{G}_1\boldsymbol{G}_4$ 上的点 $\boldsymbol{Q}$.

每个面 W_K 上的动力学是双稳的, 顶点 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_3$ 是吸引子, 如图 5.5 所示. 因此, 依赖于初值, 群体或者趋于亲社会策略, 或者趋于利己策略. 在边 $x_2=x_4=0$ 上, 亲社会策略 $\boldsymbol{G}_1$ 为风险占优当且仅当 $h+\mu(h-l)<1/2$.

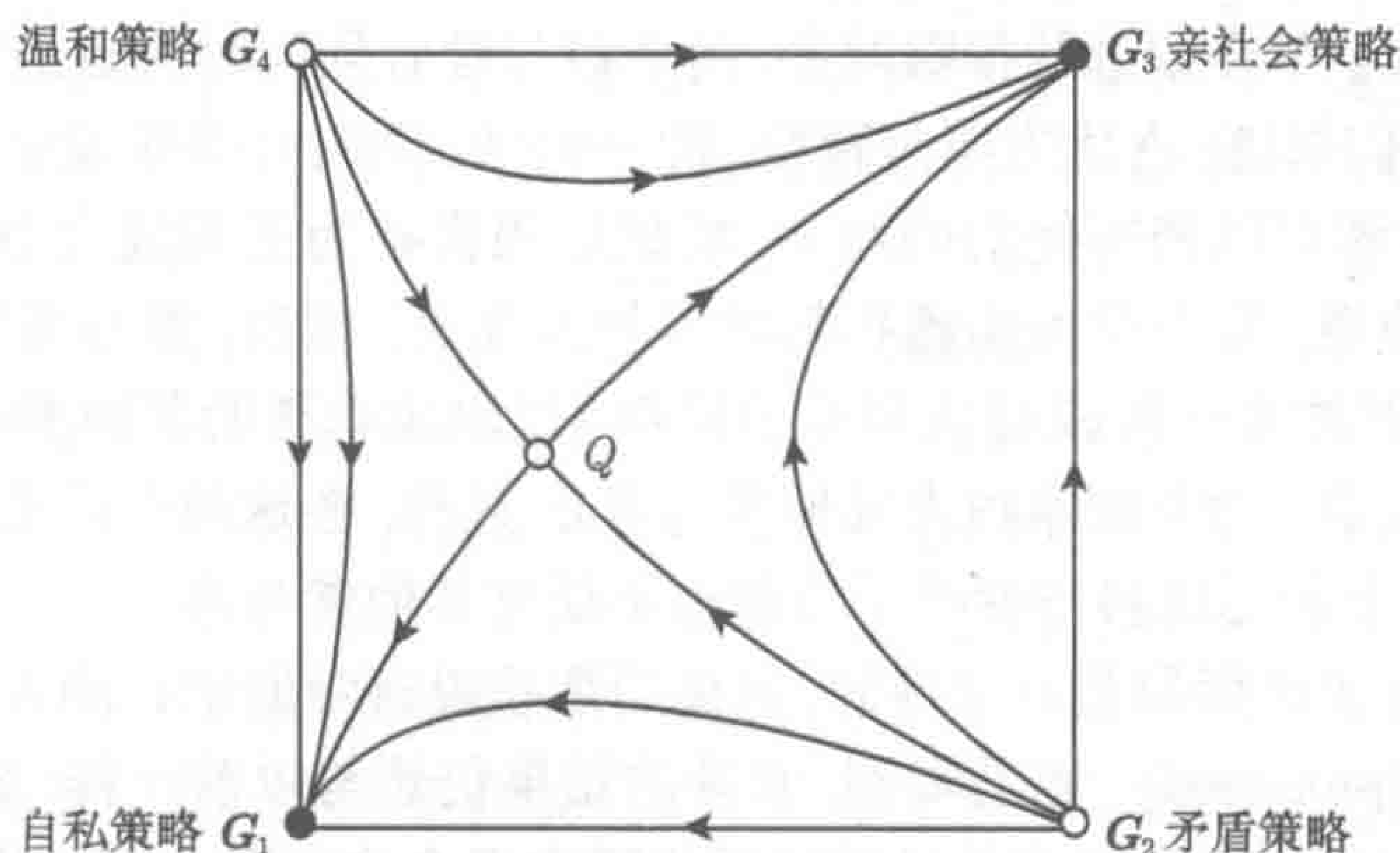


图 5.5 如果提议者知道响应者在以往博弈中所接受的报价, 简化的最后通牒博弈在 W_K 面上的复制动力学. 动力学性态是双稳的, 利己和亲社会状态都是吸引子

这表明如果考虑声誉 (和少许的利己), 使用公平策略的群体不会被侵入. 原因是, 温和策略现在处于不利地位. 同样地, 我们可以在完全最后通牒博弈中考虑基

①译注: 原文此处为 "$r>0$".

于个体的模拟. 此时假设参与者已知先前的博弈, 且提议者提供得更少: 提供的或者是 p 值, 或者是已知的响应者在前面博弈中所接受的最低值. 进化首先指向很小的 (p,q) 值, 然后沿着对角线 $p=q$ 慢慢向上攀爬. 在参数值的较大范围内, 报价趋于 40% 到 50% 之间, 如图 5.6 所示.

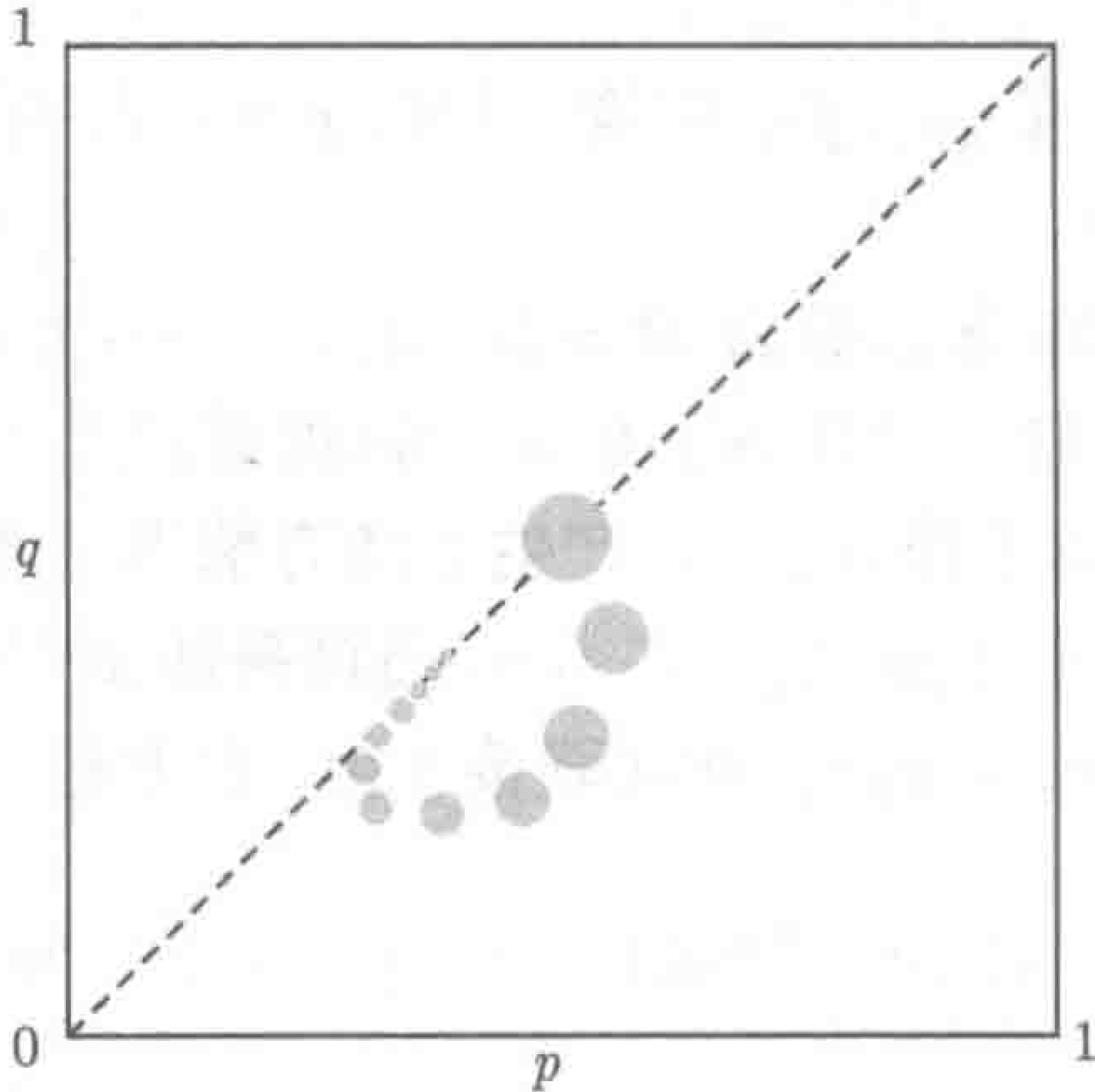


图 5.6 参与者以一个确定的概率知道彼此先前的情况, 且在不被拒绝的情况下降低报价. 在基于个体的最后通牒博弈模拟中, 报价和期望水平的进化. p 和 q 都趋于 50% 偏下的值

5.5 捐赠与威慑

现在回到 3.1 节介绍过的捐赠博弈. 两个参与者必须同时决定是否付出 c 使得对方得到 b. 我们知道, 占优方案是背叛, 即不捐赠. 现在对博弈增加第二阶段. 在第二阶段, 参与者可以损害他们的对手. 将假设损害行为是实施大小为 β 的罚款. 罚款由实验者收取, 而不是由实施罚款的参与者收取. 相反, 参与者还必须支付费用 γ. 因此, 博弈的第一阶段提供利他的机会 (付出成本帮助其他参与者), 第二阶段提供损他的机会 (付出成本损害其他参与者). 显然, 在这两个阶段, 占优方案都是避免付出. 一个利己的参与者既不帮助也不损害其他参与者.

假设参与者实施惩罚是有条件的, 只惩罚那些没有帮助他们的人. 这意味着背叛者会被惩罚 (punished). 容易看到, 长期的结果仍然与以前一样: 没有亲社会行为出现. 确实, 用 $\boldsymbol{e}_1$ 标记那些通过向对手捐赠实现合作的参与者, 用 $\boldsymbol{e}_2$ 标记那些不合作的参与者, 即那些背叛者; 类似地, 用 $\boldsymbol{f}_1$ 标记那些惩罚背叛者的参与者, 用 $\boldsymbol{f}_2$ 标记那些不这样做的参与者. 收益矩阵如下

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (-c,b) & (-c,b) \\ \boldsymbol{e}_2 & (-\beta,-\gamma) & (0,0) \end{array}. \tag{5.17}$$

虽然现在考虑的是两阶段而非两角色博弈, 但是由于情形是类似的, 所以我们还是像两角色博弈一样使用相同的记号. 尽管有所不同, 我们还是可以采用以前的方法. 确实, 两阶段博弈的每一个策略必须指定第一阶段做什么, 第二阶段做什么. 因此, 它由 $e_i\boldsymbol{f}_j$ 对给出 $(i,j\in\{1,2\})$. 如同 5.2 节, 用 $\boldsymbol{G}_1=\boldsymbol{e}_1\boldsymbol{f}_1,\boldsymbol{G}_2=\boldsymbol{e}_2\boldsymbol{f}_1,\boldsymbol{G}_3=\boldsymbol{e}_2\boldsymbol{f}_2$ 和 $\boldsymbol{G}_4=\boldsymbol{e}_1\boldsymbol{f}_2$ 分别表示四个策略. 策略 $\boldsymbol{G}_1$ 对应亲社会策略: 提供帮助, 且惩罚那些不提供帮助的人. $\boldsymbol{G}_3$ 是避免任何付出的利己策略: 既不帮助对手, 也不期望得到任何帮助. $\boldsymbol{G}_2$ 同样被看成是矛盾的: $\boldsymbol{G}_2$ 参与者选择背叛, 但是又惩罚背叛的对手. 最后, $\boldsymbol{G}_4$ 再一次被看成是温和策略: $\boldsymbol{G}_4$ 提供帮助但是不会因为没有得到回报而惩罚对手.

采用与 5.2 节相同的方法, 可以得到形如 (5.5) 的矩阵, 其中 $R=c-\beta$, $S=c$, $r=0$ 而 $s=\gamma$. 同样地, 流形 $W_K=\{\boldsymbol{x}\in S_4:x_1x_3=Kx_2x_4\}$ 不变 (当 $K>0$) 且动力学行为如图 5.3 所示. 事实上, 简化的最后通牒博弈是当 $l=\gamma,\beta=1-l$ 且 $b=c=h-l$ 时的一种特殊情形. 直观地, 在简化的最后通牒博弈中, 捐赠意味着提供高报价而非低报价. 接受者 (即响应者) 的获利 $h-l$ 等于捐赠者 (即提议者) 的付出. 惩罚是拒绝对手的报价. 惩罚使响应者付出 l, 而使得提议者失去了 $1-l$. 如果报价更低的话, 惩罚可能会更大.

在 S_4 的内部 (准确地说, 当 $x_2>0$ 或 $x_3>0$ 时) 有 $(\boldsymbol{Mx})_4>(\boldsymbol{Mx})_1$, 因此 x_4/x_1 递增, 类似的 x_3/x_2 也递增. 所以 S_4 的内部没有驻点. 这样 W_K 上的驻点是顶点 $\boldsymbol{G}_i$ 和边界 $\boldsymbol{G}_1\boldsymbol{G}_4$ 上的点. $\boldsymbol{G}_3$ 是纳什均衡而 $\boldsymbol{G}_2$ 不是. 在边界 $\boldsymbol{G}_1\boldsymbol{G}_4$ 上, $\bar{\boldsymbol{M}}=x_1(\boldsymbol{Mx})_1+(1-x_1)(\boldsymbol{Mx})_4$ 为零 (因为 $(\boldsymbol{Mx})_1=(\boldsymbol{Mx})_4=0$). 因此, 当 $(\boldsymbol{Mx})_3\leqslant 0$ 时, 即当 $x_1\geqslant c/\beta$ 时, 此边界上的点 $\boldsymbol{x}$ 是饱和的. (条件 $(\boldsymbol{Mx})_2\leqslant\bar{M}$ 约化成相同的不等式.) 如果 $c>\beta$, 则 $\boldsymbol{G}_3$ 是唯一的纳什均衡. 这种情形不太有趣.

从现在开始, 我们将注意力集中在 $c<\beta$ 的情形: 惩罚的代价大于捐赠. 用 $\boldsymbol{Q}$ 表示 $(c/\beta,0,0,(\beta-c)/\beta)$, 闭线段 $\boldsymbol{QG}_1$ 由纳什均衡组成.

在边界 $\boldsymbol{G}_2\boldsymbol{G}_4$ 上, 还有一个驻点 $(0,c/(\beta+\gamma),0,(\beta+\gamma-c)/(\beta+\gamma))$. 它在边界以及面 $x_3=0$ 上是吸引的, 但在面 $x_1=0$ 上排斥. 边界 $\boldsymbol{G}_1\boldsymbol{G}_3$ 上也有一个驻点 $\boldsymbol{P}=((c+\gamma)/(\beta+\gamma),0,(\beta-c)/(\beta+\gamma),0)$. 它在面 $x_4=0$ 上是吸引的, 但在面 $x_2=0$ 上排斥. 在没有其他策略时, 策略 $\boldsymbol{G}_1$ 和 $\boldsymbol{G}_3$ 是双稳的. 策略 $\boldsymbol{G}_1$ 风险占优 (即它有更大的吸引盆) 当且仅当 $2c<\beta-\gamma$. 注意到在最后通牒的迷你博弈的特殊情形中, 此条件约化为 $h<1/2$.

除了 $\boldsymbol{G}_3$ 和线段 $\boldsymbol{QG}_1$, 没有其他纳什均衡. 依赖初值条件, S_4 内部的轨道不是趋于 $\boldsymbol{G}_3$ 就是趋于 $\boldsymbol{QG}_1$ 上的纳什均衡. 在只由 $\boldsymbol{G}_1$ 和 $\boldsymbol{G}_4$ 组成的群体中, 没有优势策略. 假设状态 $\boldsymbol{x}$ 通过中性漂变 (反映了系统的随机冲击) 沿边界波动. 偶尔地, 随机冲击将引入少量的 $\boldsymbol{G}_2$ 或 $\boldsymbol{G}_3$ 策略. 如果此时 $\boldsymbol{x}$ 在 $\boldsymbol{QG}_1$ 上, 选择会使状态回到边界上, 但更接近于 $\boldsymbol{Q}$ (因为 x_4/x_1 增加). 一旦状态到达 $\boldsymbol{QG}_4$, 且少量 $\boldsymbol{G}_3$ 被引

入, G_3 会受到选择的青睐最终在群体中固定下来. 利己状态 G_3 最终会胜出.

5.6 声誉的威慑作用

假设合作者 (e_1 参与者) 背叛非惩罚者 (f_2 参与者) 的概率是 μ. (因此, μ 是① f_2 类型暴露且② e_1 类型决定背叛的概率.) 类似的, 假设 ν 为背叛者 (e_2 参与者) 与惩罚者 (f_1 参与者) 合作的概率. (因此, ν 是① f_1 类型暴露且② e_2 类型决定捐赠的概率.) 则收益矩阵变为

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (-c,b) & (-c(1-\mu),b(1-\mu)) \\ \boldsymbol{e}_2 & (-(1-\nu)\beta-\nu c,-(1-\nu)\gamma+\nu b) & (0,0) \end{array}. \tag{5.18}$$

这样, $R=(1-\nu)(c-\beta)<0, S=c(1-\mu)>0, s=\gamma-\nu(b+\gamma)$ 且 $r=-b\mu<0$. 因此, 边界 G_1G_4 不再由驻点构成, 而是一条趋于 G_1 的轨道. 动力学行为如图 5.5 所示. 参数 μ 很重要 (μ 为亲社会参与者在免受惩罚的情况下选择背叛的概率). 相比之下, 只要 ν 很小, 它对动力学行为不会产生影响. 因此, 本节中取 $\nu=0$.

这样, 在 S_4 的内部存在一条由驻点 $\boldsymbol{x}$ 组成的线, 即 $x_i=m_i+\xi$, $i=1, 3$, 和 $x_i=m_i-\xi$, $i=2, 4$, 其中 ξ 是参数而

$$\boldsymbol{m}=\frac{1}{(\gamma+b\mu)(\beta-c\mu)}(c\gamma(1-\mu),bc\mu(1-\mu),b\mu(\beta-c),\gamma(\beta-c)). \tag{5.19}$$

正如简化的最后通牒博弈, 这条驻点线穿过四边形 $G_1G_2G_3G_4$ 且与每一个 W_K 面恰好交于一点. 由于 $Rr>0$, 在相应的 W_K 面上此点为复制系统的一个鞍点, 如图 5.5 所示. 在每一个面, 从而在 S_4 内部的动力学行为是双稳的, 吸引子为 G_1 和 G_3. 依赖于初值, 除了零测集以外, 每一个轨道都趋于两个状态中的一个.

再一次, 当 $\mu\to 0$ 时, 点 $\boldsymbol{m}$ 和所有内部驻点趋于点 $\boldsymbol{Q}$. $\mu=0$ 存在一个高度退化分支. 这条 (非常短的) 不动点线段被不动点的横截线即 G_1G_4 所代替, 其中一段 QG_1 由纳什均衡构成.

因此, 引入任意小的扰动 μ 将改变群体的长期行为. 动力学系统不再趋于利己状态 G_3 (背叛, 不惩罚), 而是具有两个吸引子, 即 G_3 和亲社会状态 G_1 (合作, 惩罚背叛者). 注意, μ 正比于获知对手类型的概率.

现在直接考虑 $\mu=1$, 这表示完全知道对手的类型. 此时 $S=0$. 这就产生了 $\mu=0$ 的镜像情形. 现在 G_3G_4 边界由驻点构成, 在 $\hat{Q}G_3$ 上的点均为纳什均衡, 其中 $\hat{\boldsymbol{Q}}=(0,0,b/(b+\gamma),\gamma/(b+\gamma))$, 而扰动导致最终状态趋于另一个纳什均衡, 即 G_1. 在完全信息下, 亲社会参与者如果在不受惩罚的情况下总选择背叛, 那么他们必占上风.

5.7 揭示错误

先前的模型在某种意义上并不完整, 它完全依赖于引入声誉来改变边界 G_1G_4 上的动力学行为. 但在那条边上, 群体只由两种类型组成且在博弈的捐赠阶段都选择合作. 假设经过若干回合都没有背叛者出现, 参与者如何辨别对手是 $\boldsymbol{f}_1$ 还是 $\boldsymbol{f}_2$, 即对手是否愿意去惩罚背叛者?

有几种途径可以解决这个问题. 一是假设参与者从其他途径去了解对手是否有惩罚倾向. 确实, 通过观察对手同吵闹的孩子互动或者观察他们对每日新闻的评论, 而不是仅仅通过观察他们在捐赠博弈中如何行动来判断他们是脾气暴躁还是性格温顺似乎是可能的.

但最简单的方法是引入误差. 假设每个参与者反复地参与博弈 (当然, 永远不会遇到同一个对手), 且参与者想要捐赠却有 ϵ 的概率未能实施. (这可能是由于失误, 或者资源的短暂缺失所造成.) 在缺乏声誉时, 有如下的收益结构:

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (-(1-\epsilon)c-\epsilon\beta,(1-\epsilon)b-\epsilon\gamma) & (-(1-\epsilon)c,(1-\epsilon)b) \\ \boldsymbol{e}_2 & (-\beta,-\gamma) & (0,0) \end{array}. \tag{5.20}$$

与 5.5 节的情况相比, s 仍未改变, 而 R 和 S 都乘以了 $(1-\epsilon)$, 没有改变符号. 因此, 在对应的边界保持了动力学行为. 但现在 r 等于 $\epsilon\gamma$, 因而符号为正. 这意味着在边界 G_1G_4 上, 解流指向 G_4: 惩罚被占优. 即图 5.2 中情形 (a) 的动力学行为. 单形 S_4 内部的所有轨道都趋于顶点 G_3, 利己类型获胜.

现在引进声誉. 为简化起见, 假设参与者在知道对手不是惩罚者的情况下, 决不捐赠, 即总是背叛. (仅假设参与者以小概率背叛即可.) 这样, 参数 μ 即为知道对手在过去未能惩罚背叛者的概率. 如果有完全信息, 这只会发生在对手是 $\boldsymbol{f}_2$ 类型且至少遇到过一次背叛的情形. 在边 $x_2=x_3=0$ 上, 所有的参与者愿意捐赠, 且只会因为失误而发生背叛. 对手经历 k 回合而没有遇见失误的背叛其概率是 $(1-\epsilon)^k$. 如果回合数呈几何分布, 且下一轮概率是常数 $w<1$, 那么对手经历 k 回合的概率是 $w^k(1-w)$. 这意味着

$$\mu=\frac{w\epsilon}{1-w(1-\epsilon)}. \tag{5.21}$$

如果假设 $\boldsymbol{e}_1$ 参与者因失误或者知道对手是 $\boldsymbol{f}_2$ 类型而背叛, 则得到

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (-(1-\epsilon)c-\epsilon\beta,(1-\epsilon)b-\epsilon\gamma) & (-(1-\epsilon)(1-\mu)c,(1-\epsilon)(1-\mu)b) \\ \boldsymbol{e}_2 & (-\beta,-\gamma) & (0,0) \end{array}. \tag{5.22}$$

这样 $r = \epsilon\gamma - \mu(1-\epsilon)b$ 为负, 如果

$$\gamma < \frac{w(1-\epsilon)b}{1-w(1-\epsilon)}, \tag{5.23}$$

即用于惩罚背叛者的费用不是很高. 此时, 动力学行为如图 5.5 所示, 亲社会策略 G_1 为吸引子.

当然, 这也可应用于最后通牒博弈. 此时, $r = \epsilon\gamma - \mu(1-\epsilon)b$ 为负, 如果

$$l < w(1-\epsilon)h, \tag{5.24}$$

即低报价较高报价足够的小.

5.8 信任博弈

信任博弈是一个两参与者博弈, 与最后通牒博弈形成了巧妙的对应. 首先, 由抛硬币决定两个参与者谁是提议者即投资者. 然后投资者向响应者或者说受托人捐赠固定金额 c, 且投资者知道金额会被乘以一个因子 $r>1$. 然后, 受托者可以选择还给投资者总金额的一部分 β (返还的金额不会被乘以因子而增加). 博弈到此结束.

显然, 一个利己的受托者将不返还任何金额. 如果投资者知道会这样, 就不会提供任何金额. 在真实的实验中, 投资者通常会捐赠, 受托者通常也会返还足够的金额使双方都盈利.

在简化的信任博弈中, 假设金额 c 和 β 是固定的. 投资者只需要决定是否给予受托者金额 c. 因此投资者只有两种选择 $\boldsymbol{e}_1$ (捐赠) 和 $\boldsymbol{e}_2$ (背叛). 类似的, 接受捐赠 (即 $b=rc$) 的受托者也有两种选择, 返回金额 β 或者不返回: 这两种选择将分别用 $\boldsymbol{f}_1$ 和 $\boldsymbol{f}_2$ 表示. 为了让博弈更加有趣, 假设 $c<\beta<b$. 此时, 如果参与者双方选择合作, 则都能得利. 此博弈的收益矩阵为

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (\beta-c, b-\beta) & (-c, b) \\ \boldsymbol{e}_2 & (0,0) & (0,0) \end{array}. \tag{5.25}$$

因为两个参与者有相同的概率成为投资者或受托者, 所以实际上处于一个对称博弈之中. 在分析之前, 先考虑另一个稍微一般的博弈, 与 5.5 节研究过的具有惩罚背叛机制的捐赠博弈相比, 它展示了一个有趣的互补性. 这种具有奖励捐赠机制的捐赠博弈也是一个两阶段博弈. 第一阶段就是 3.1 节描述的捐赠博弈. 第二阶段, 接受者可以选择返还捐赠的一部分. 假设这将花费 γ, 而产出 β 奖励对手 (如果 $\beta=\gamma$,

这是简单的返回). 现在假设 $0 < c < \beta$ 且 $0 < \gamma < b$. 如果 e_1 和 e_2 是第一阶段的两个选择 (捐赠或背叛), f_1 和 f_2 是第二阶段的两个选择 (奖励捐赠或不奖励), 那么收益结构为

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (\beta - c, b - \gamma) & (-c, b) \\ \boldsymbol{e}_2 & (0,0) & (0,0) \end{array} . \tag{5.26}$$

简化的信任博弈可以看成是此博弈的特殊情况 (如果考虑两角色博弈和两阶段博弈之间的对应关系). 这里有四个策略, 也就是 (a) 亲社会策略 $\boldsymbol{G}_1 = \boldsymbol{e}_1\boldsymbol{f}_1$ (捐赠, 奖励); (b) 不信任的策略 $\boldsymbol{G}_2 = \boldsymbol{e}_2\boldsymbol{f}_1$ (背叛, 但奖励捐赠者); (c) 利己策略 $\boldsymbol{G}_3 = \boldsymbol{e}_2\boldsymbol{f}_2$ (不捐赠, 不奖励); 最后, (d) 策略 $\boldsymbol{G}_4 = \boldsymbol{e}_1\boldsymbol{f}_2$ (捐赠, 但不奖励). 对应于收益矩阵 (5.5), 有 $R = c - \beta < 0, r = \gamma > 0, S = c > 0$ 和 $s = 0$, 如图 5.7 所示.

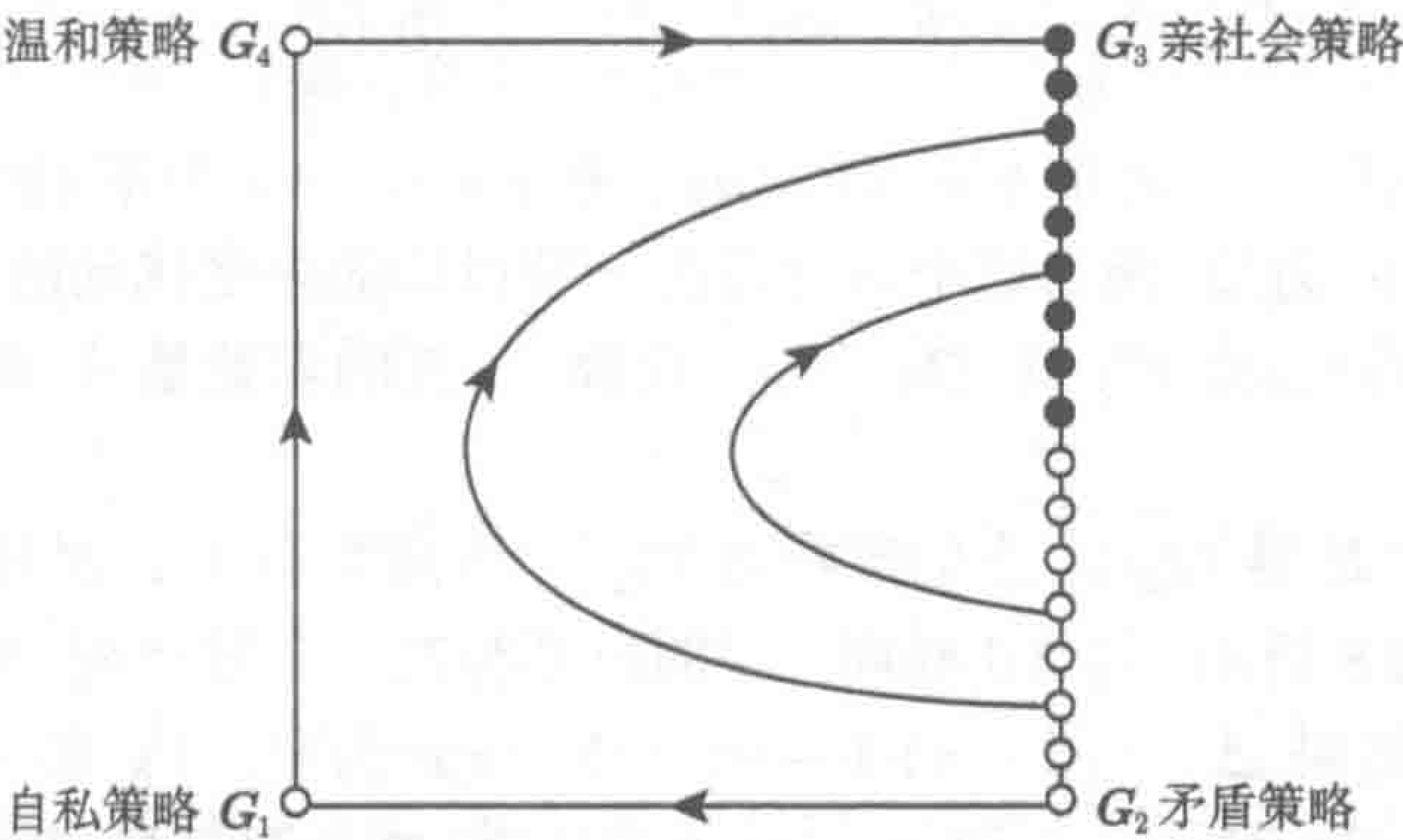


图 5.7 简化的信任博弈的复制动力学. 边 $\boldsymbol{G}_2\boldsymbol{G}_3$ 由驻点组成. 从长远来看, 没有人信任对手

如果 $x_3 = x_4 = 0$, 即如果群体中每个参与者都愿意奖励捐赠者, 那么最好捐赠, 即 $\boldsymbol{G}_1$ 对 $\boldsymbol{G}_2$ 占优. 如果 $x_2 = x_3 = 0$, 即捐赠是理所当然的, 那么最好不奖励, 即 $\boldsymbol{G}_4$ 对 $\boldsymbol{G}_1$ 占优. 如果 $x_1 = x_2 = 0$, 即没有参与者奖励捐赠, 那么 $\boldsymbol{G}_3$ 对 $\boldsymbol{G}_4$ 占优, 即最好不捐赠. 最后, 如果 $x_1 = x_4 = 0$, 即没有参与者捐赠, 那么无所谓奖不奖励. 此时, 群体的每一个状态都是驻点. $\boldsymbol{G}_2$ 和 $\boldsymbol{G}_3$ 都没有优势.

容易看到, 当

$$\boldsymbol{Q} = \left(0, \frac{c}{\beta}, \frac{\beta - c}{\beta}, 0\right) \tag{5.27}$$

时, 线段 $\boldsymbol{Q}\boldsymbol{G}_3$ 由饱和驻点, 即纳什均衡组成. 实际上, 当 $x_1 = x_4 = 0$ 时, $(\boldsymbol{M}\boldsymbol{x})_1$ (规范化为零) 和 $(\boldsymbol{M}\boldsymbol{x})_4$ 都比平均收益 $\bar{\boldsymbol{M}} = (\boldsymbol{M}\boldsymbol{x})_2 = (\boldsymbol{M}\boldsymbol{x})_3 = c - \beta x_2$ 小. 对比简化的最后通牒博弈, 解流沿着边界从 $\boldsymbol{G}_2$ 到 $\boldsymbol{G}_1$ 再到 $\boldsymbol{G}_4$, 最后到 $\boldsymbol{G}_3$. 内部的所有轨道当 $t \to +\infty$ 时, 趋于线段 $\boldsymbol{Q}\boldsymbol{G}_3$, 而当 $t \to -\infty$ 时, 趋于线段 $\boldsymbol{Q}\boldsymbol{G}_2$. 因此, 长期下去, 群体将只由总是背叛 (因此没有奖励) 的参与者组成. 特别地, 简化的信任

博弈永远不会有所进步: 没有捐赠, 没有回报.

5.9 报酬与声誉

现在引入声誉的影响. 假设合作者 ($\boldsymbol{e}_1$ 类型参与者) 知道其对手不会提供奖励 ($\boldsymbol{f}_2$ 类型参与者) 时, 他们背叛的概率为 μ. 因此, μ 是① $\boldsymbol{f}_2$ 类型暴露且② $\boldsymbol{e}_1$ 类型决定背叛的概率. 类似的, 假设背叛者 ($\boldsymbol{e}_2$ 类型参与者) 知道他们有可能得到奖励时选择合作的概率为 ν. 因此, ν 是① $\boldsymbol{f}_1$ 类型暴露且② $\boldsymbol{e}_2$ 类型做出相应反应而捐赠的概率. 这给出了收益结构

$$\begin{array}{c|cc} & \boldsymbol{f}_1 & \boldsymbol{f}_2 \\ \hline \boldsymbol{e}_1 & (\beta - c, b - \gamma) & (-c(1-\mu), b(1-\mu)) \\ \boldsymbol{e}_2 & ((\beta - c)\nu, (b-\gamma)\nu) & (0,0) \end{array} . \tag{5.28}$$

现在, $R = (c - \beta)(1 - \nu) < 0$, $S = c(1 - \mu) > 0$, $r = \gamma - b\mu$ 为正 (如果 μ 足够小的话) 而 $s = (\gamma - b)\nu$ 为负. 最后这个条件与前一节讨论的未受扰动的系统不同. 边界 $\boldsymbol{G}_2\boldsymbol{G}_3$ 不再由驻点构成. $\boldsymbol{G}_2$ 对 $\boldsymbol{G}_3$ 占优. 这样, 本质的参数是 ν, 而下面的讨论中取 $\mu = 0$.

当 $\nu > 0$ 时, 边界 $\boldsymbol{G}_2\boldsymbol{G}_3$ 上的流指向 $\boldsymbol{G}_2$①, 张成鞍面 W_K 的框架上的流是循环保向的, 如图 5.8 所示. 与 5.6 相同, S_4 的内部存在一条驻点线. 可以证明 W_1 面由周期轨构成. 如果 $\Delta := (\beta - \gamma)(1 - \nu) - (b - c)\nu$ 为负, W_K $(0 < K < 1)$ 上的所有非平衡轨道 (以及 $x_1 = 0$ 和 $x_3 = 0$ 上的轨道) 螺旋离开驻点线且趋于异宿环 $\boldsymbol{G}_1\boldsymbol{G}_4\boldsymbol{G}_3\boldsymbol{G}_2$. W_K $(K > 1)$ 上的所有非平衡轨道 (与面 $x_2 = 0$ 和 $x_4 = 0$ 上的轨道一样) 螺旋离开异宿环且趋于驻点线. 如果 Δ 为正, 则流向相反.

需要强调的是当 $\nu > 0$ 且 $\Delta \neq 0$ 时的高度不可预测的动力学行为. 鞍面 W_1 将状态空间 S_4 分成两个大小相等的区域. 当初值条件在其中一个区域时, 复制系统使状态趋于驻点线. 但随机波动最终会导致状态进入另一个区域, 此时复制系统将使轨道趋于异宿环 $\boldsymbol{G}_1\boldsymbol{G}_4\boldsymbol{G}_3\boldsymbol{G}_2$. 群体似乎长时间处于一种策略状态, 然后突然切换到下一种策略状态, 且仍然停留很长的一段时间, 然后循环下去. 然而, 任意小的随机冲击可能使得状态回到前面那个区域, 且动力系统使流再一次趋于驻点线, 等等. 此时策略频度的时间平均不收敛. 只能说, 最可能的群体状态要么是单态 (即接近 S_4 的一个角), 要么是接近驻点线 (四种策略类型均非零. 当 ν 很小时, 奖励者的频度接近 c/β, 且捐赠者的频度很小) 的吸引部分.

因此, 在进一步的经济行为中, 正激励的效率似乎大大低于负激励的效率. 需要强调的是, 这适用于基于随机配对的模型. 如果参与者可以主动选择合作伙伴,

①译注: 原文此处为 "$\boldsymbol{G}_3$".

那么拥有奖励合作者的声誉比拥有惩罚背叛者的声誉更受欢迎.

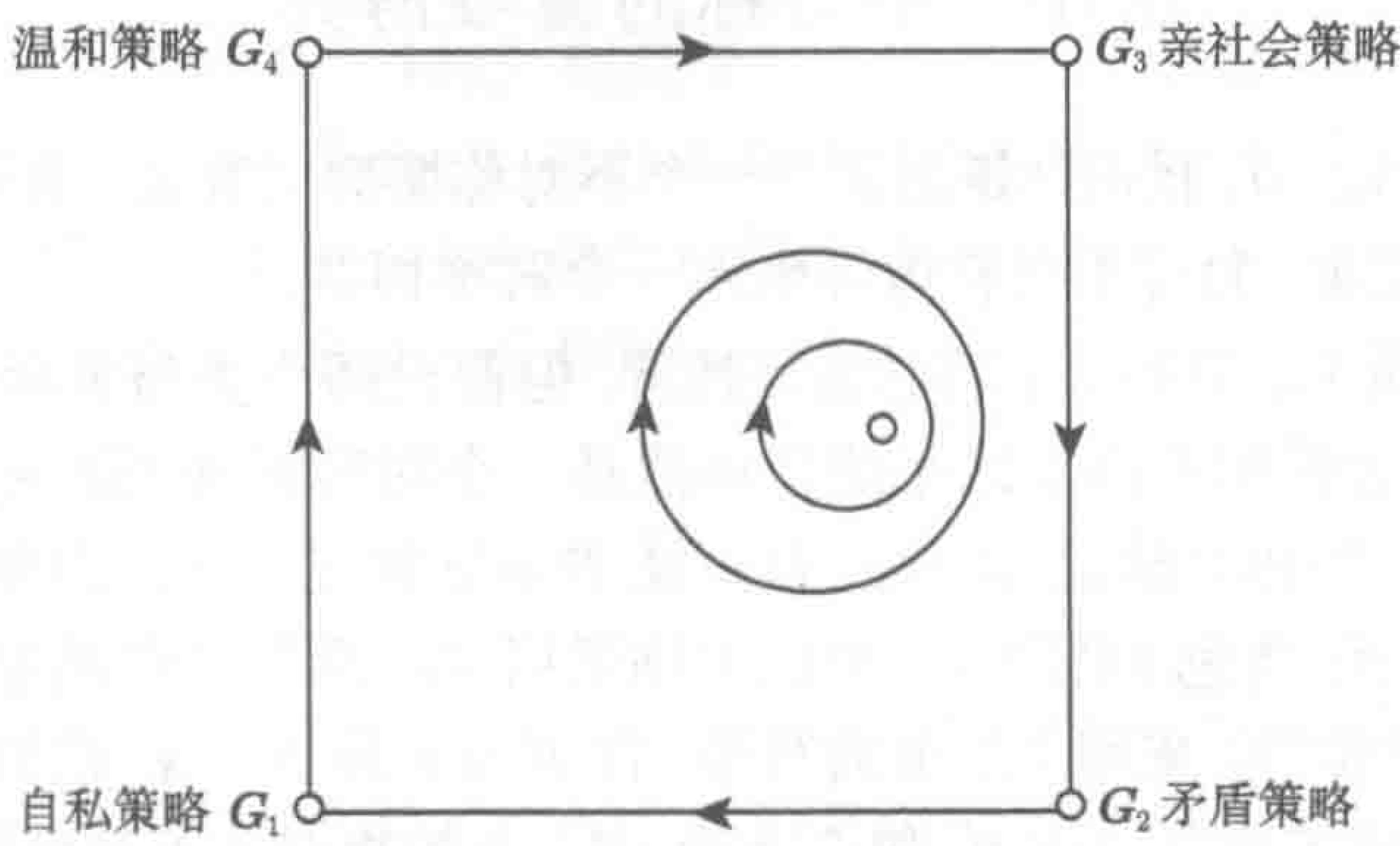


图 5.8 简化信任博弈的复制动力学, 接受者通过奖励捐赠者而获得声誉, 依赖于比率 x_1x_3/x_2x_4 的初值, 轨道或者螺旋趋于内部驻点, 或者趋于四条边构成的环. 如果比率为 1, 则轨道都是闭轨

这样就面临与 5.7 节中最后通牒博弈相同的问题. 当 $x_1 = x_4 = 0$ 时, 无人捐赠. 此时, 奖励捐赠的特征是如何出现的呢? 像以前那样假设参与者偶尔失误并不自然, 原因是因失误而产生捐赠比在试图捐赠中出现失误更不可能.

最后, 考虑参与者拒绝奖励的行为被公开会怎样. 此时, 这样的参与者不太可能得到捐赠. 这意味着 μ 接近于 1, 参数 $r = \gamma - b\mu$ 为负. 这样, 所有内部的轨道趋于亲社会状态 G_1, 如图 5.9 所示. 这特别适用于银行和基金的信任博弈. 假设基金经理不对投资进行返还会被公开, 则亲社会策略 (作为客户选择投资, 作为基金经理返还多于投资的金额) 是全局吸引子.

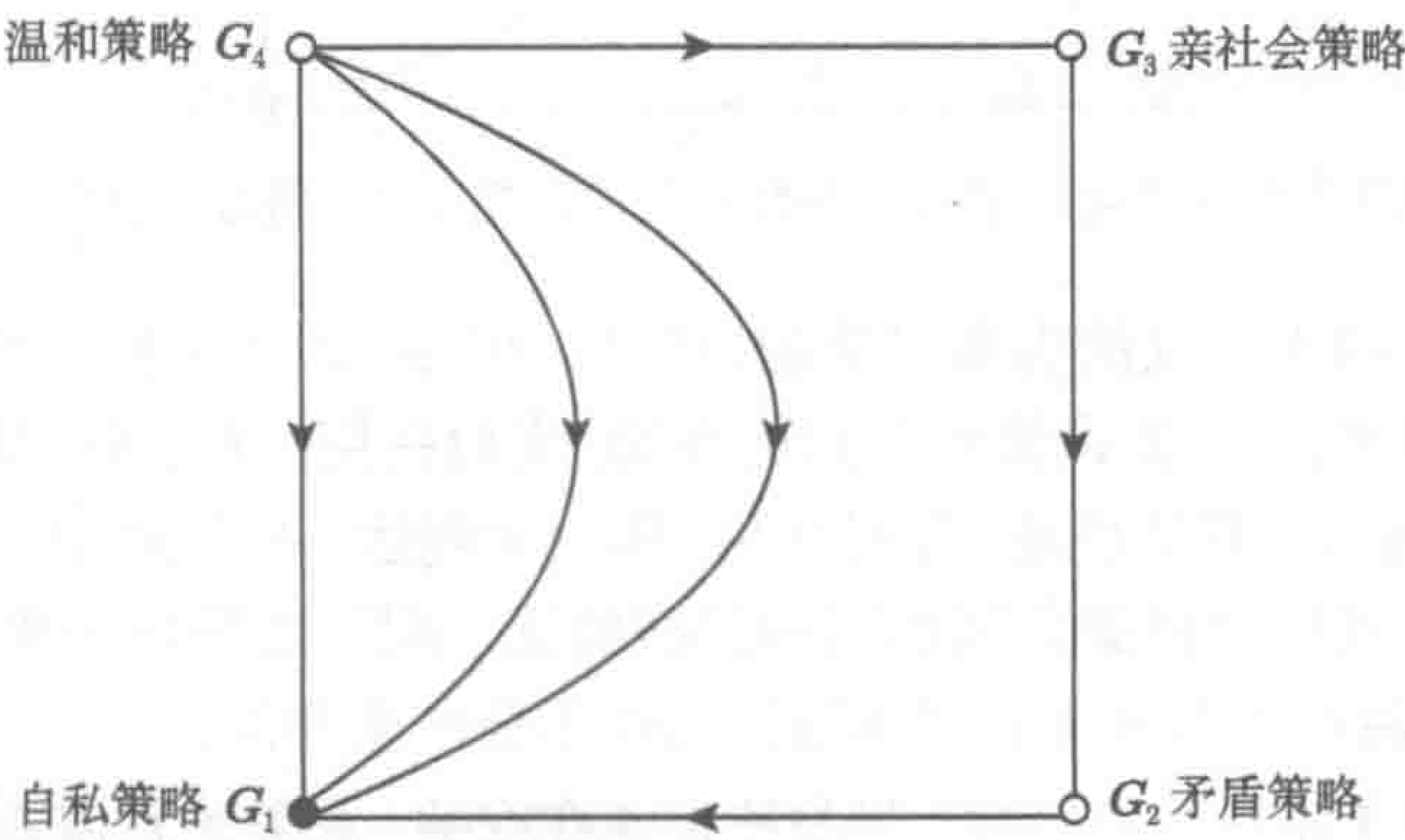


图 5.9 在参与者不返还投资可能声名狼藉的情况下, 简化信任博弈的复制动力学. 亲社会状态 G_1 吸引所有轨道

5.10 不对称的雪堆博弈

在 2.5 节和 5.2 节, 我们知道怎样将一个不对称博弈对称化. 有时, 考虑相反的方案是有用的. 比如, 允许不对称博弈嵌入一个对称博弈.

特别地, 回到 1.4 节和 3.1 节的雪堆博弈, 但假设两个参与者存在差异. (例如, 他们年龄不同或者性别不同.) 这个博弈仍然是一个对称博弈 (即 $\boldsymbol{e}_i = \boldsymbol{f}_i, i = 1, 2$, 收益结构由 (5.2) 给出且满足 $A = a$, $D = d$, $B = c$ 和 $C = b$), 但参与者可以使用有条件的策略类型: 角色 I采用 $\boldsymbol{e}_1$, 角色 II则采用 $\boldsymbol{f}_2$. 尽管角色的差异并不影响收益, 但它可以产生惯性. 采用 5.2 节的符号, 有 $R = r$ 且 $S = s$. 在雪堆博弈的背景下, $R = c/2 > 0$ 且 $S = c - b < 0$. 每个流形 W_K 上的相图关于对角线是对称的, 且是双稳的. 策略 $\boldsymbol{G}_2$ 和 $\boldsymbol{G}_4$ 是吸引子. 对于不同的角色, 相图给出了对应的行为. 几乎对所有的初始条件, 状态收敛到两个导致合作的条件策略之一: 参与者是否付费依赖于他们扮演的角色 (图 5.10).

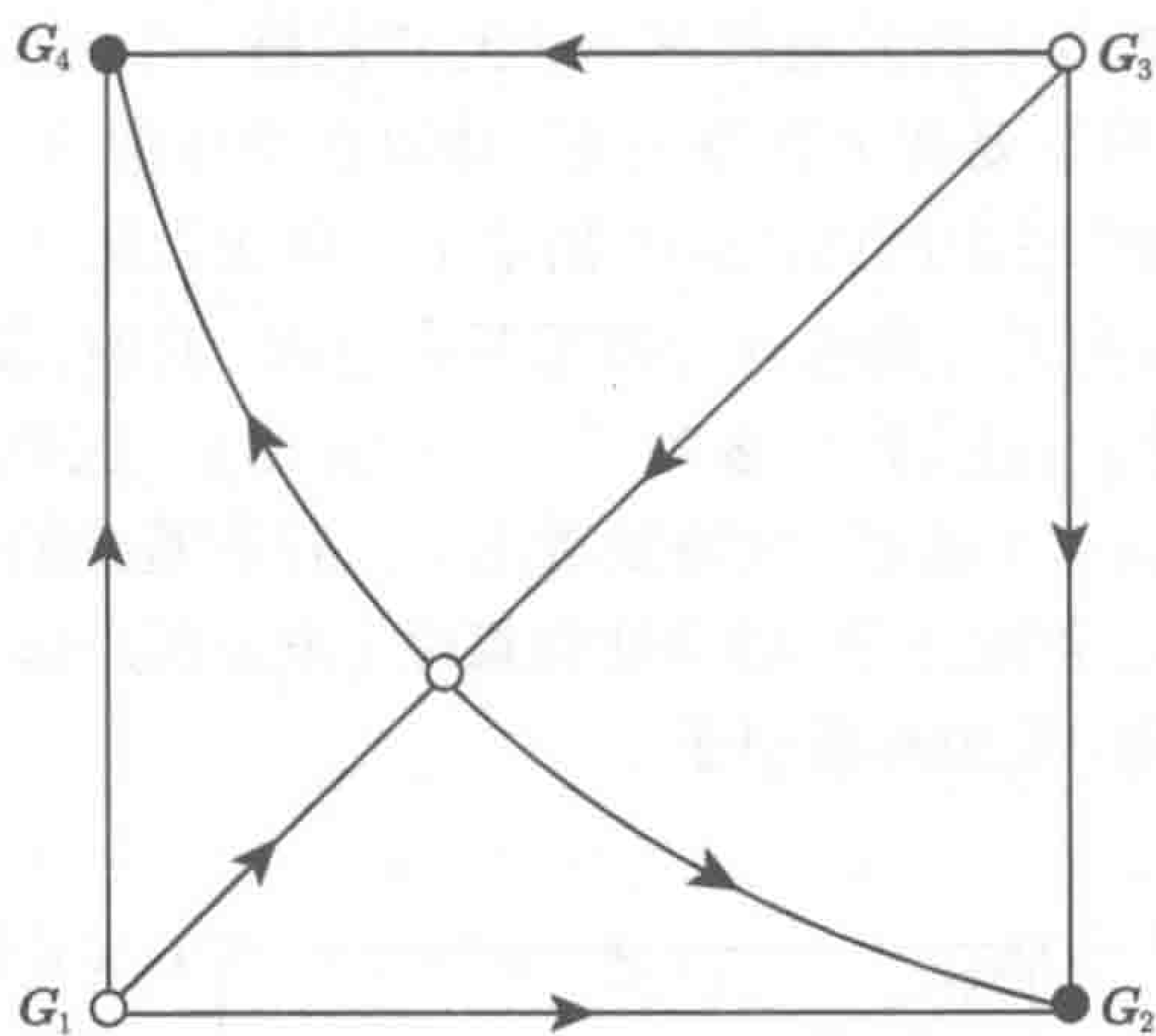


图 5.10 在两个参与者可以通过一些特征区别的情况下, 雪堆博弈的复制动力学

这意味着先前无关联的因素可能确定结果. 例如, 考虑年龄或到达顺序. 条件策略可以是这种形式: “如果第一个到达现场, 支付费用; 第二个到达, 不用支付.” 对称的: “如果第二个到达现场, 支付费用; 第一个到达, 不用支付”. 如果通过外部影响利用某种方式将两结果之间的对称打破的话. 则不对称所产生的动力学行为具有稳定性. 随后的不对称 (如, 性别或大小) 不会产生作用.

对两策略博弈而言, “去对称” 只与稳定共存有关. 如果对称博弈导致占优或双稳, 那么引入两角色不会影响结果: 对应的不对称博弈仍然导致占优或双稳. 但如果对称博弈是稳定共存的, 那么引入两角色会完全改变结果: 相应的不对称博弈是

双稳的.

特别地, 我们知道在重复的雪堆博弈, 就平均而言, 两个参与者在每一轮都合作, 与他们交替合作和背叛一样好. 但是第二个选择似乎不太可能. 更可能的是, 一旦某个参与者背叛而另一个合作, 这种不对称将在后续回合中产生惯性. 此时, 一个参与者总是选择合作而另一个总是背叛与他们每轮都合作的总收益是相同的. 但一个参与者总是都被另一个参与者剥削.

5.11 参考文献

最后通牒博弈由 Güth, Schmittberger 和 Schwarze (1982) 引入, 在实验经济学中发挥了巨大的作用, 参看 Camerer (2003), Rankin (2003), Camerer 和 Fehr (2006) 以及 Henrich 等 (2006). 进化博弈的理论研究, 参看 Page, Nowak 和 Sigmund (2000), Page 和 Nowak (2001), 以及 Härdling (2007). 5.2 节的对称化博弈可以参看 Gaunersdorfer, Hofbauer 和 Sigmund (1991), 以及 Hofbauer 和 Sigmund (1998). Nowak, Page 和 Sigmund (2000) 分析了简化的最后通牒博弈. 基于个体的模拟可以查看网页`http://www.wuwien.ac.at/usr/ma/hbrandt/ultimatum/`. Sigmund, Hauert 和 Nowak (2001) 研究了惩罚机制以及奖励机制对捐赠博弈的影响, 推广到公共利益博弈参看 Hauert, Haiden 和 Sigmund (2004). Berg, Dickhaut 和 McCabe (1995) 引入了信任博弈, Güth 等 (2001), Fehr 和 Fischbacher (2003), McCabe, Rigdon 和 Smith (2003), Bohnet 和 Croson (2004), Cox (2004), 以及 Eckel 和 Wilson (2004). DeBruine (2002)① 的实验显示出信任度会因为亲缘关系而提高. Johnson 和 Bering (2006) 认为对超自然惩罚的恐惧对合作的出现起到了巨大的作用. Isaac 和 Walker (1988), 以及 Rankin (2003) 探讨了沟通的作用. 在 Gneezy 和 Rustichini (2000) 或者 Fehr 和 Rockenbach (2003) 的实验中, 显现了惩罚的负面作用. Maynard, Smith (1982), 以及 Sugden (1986) 和很多其他人研究了双稳态对称博弈的去对称作用 (通常在冲突升级的背景下).

①译注: 原文此处为 “de Bruine (2005)”.

第 6 章　公共利益和共同努力: 自由与强制之间

6.1　公共利益博弈

到目前为止, 我们只考虑了两两博弈. 但许多交互都发生在大型团体中. 这就引出新的视角. 特别地, 报答变得更加困难. 如果在一个群体中反复博弈, 一个对手背叛而另一个对手合作, 那报答谁呢?

另一方面, 在多于两个成员的群体中, 可能形成多数派, 这可以促进合作. 就像汉密尔顿 (W. D. Hamilton) 在《人与生俱来的社会性》(*Innate Social Aptitudes of Man*) 中写到的那样: "令人欣慰的是, 人生是多人游戏 (博弈, game), 而不仅仅是一些孤立的两人游戏的聚集."

现在, 团体中的交互将会被模型化为所谓的公共利益博弈. 典型的例子有群体围猎、共同建造住所或保护公共资源、相互合作以抵御来自内外的威胁.

这样的公共利益博弈显现出一个社会困境: 背叛者比合作者更得利. 引入惩罚并不能消除这个困境, 因为惩罚自身也是公共利益. 但我们将会看到, 如果参加博弈是自愿的而不是强制的, 那么基于对背叛者惩罚的合作将会出现.

6.2　模型化公共利益博弈

假设有 N ($\geqslant 2$) 个个体参加公共利益博弈. 每个参与者必须决定在博弈中选择合作还是背叛. 我们将考虑两种模型.

首先, 考虑 SR (自返还, self-return) 的情形, 假设 N_c 个贡献者的捐赠都被乘以 $r > 1$, 然后在 N 个参与者中进行平分 . 因此每个参与者从公共利益中获得的收益为

$$\frac{rcN_c}{N}. \tag{6.1}$$

此时, 贡献者必须支付固定的成本 c. 注意到, 贡献者投资额的 r/N 将返还给自己. 这个博弈只有当 $r < N$ 时才为社会困境.

第二种情形为 OO (仅对他人, others-only), 每个贡献者的捐赠被乘以 r, 然后在其他 $N-1$ 个参与者之间平分. 此时, 贡献者没有从自己的投资中得到回报. 背叛者的收益为

$$\frac{rcN_c}{N-1}. \tag{6.2}$$

贡献者从公共利益中获得的收益为

$$\frac{rc(N_c-1)}{N-1}, \tag{6.3}$$

同时还必须支付成本 c.

一个参与者从背叛变为合作将会承担成本, 在 SR 情形为 $c(1-\frac{r}{N})$, 在 OO 情形为 c. 如果所有参与者都合作, 则参与者在这两种情形获得的收益相同, 都为 $(r-1)c$.

注意到, 如果只有两个参与者, 即 $N=2$, 则 OO 情形的公共利益博弈即为 3.1 节的捐赠博弈. 问题简化为是否支付成本 c 使得对手获得收益 $b:=rc$. 当 $1<r<2$ 时, SR 情形的公共利益博弈也是一个捐赠博弈. 此时, 收益为 $rc/2$, 而成本为 $c(1-\frac{r}{2})$.

现在考虑无限大群体的极限情形, 假设具有 N 个参与者的随机样本进行公共利益博弈. 如果 x 表示贡献者的频度, y 表示背叛者的频度, 则在 OO 情形, 贡献者的期望收益为 $P_x=c(rx-1)$, 背叛者的为 $P_y=crx$. 将这两个表达式乘以 $(N-1)/N$, 就得到 SR 情形对应的收益值. 每种情形 P_y 都比 P_x 大, 所以背叛者将占上风.

这个社会困境可以通过针对个别参与者的正向或负向激励克服吗? 这里, 我们将只研究惩罚的作用.

6.3 带有惩罚的公共利益博弈

第 5 章已经研究了两两作用的带有惩罚的公共利益博弈, 即 $N=2$ 时的情形. 容易将这种分析推广到更大的 N. 博弈包括两个阶段. 第一阶段, 参与者必须在 $\boldsymbol{e}_1$ (向公共利益贡献) 与 $\boldsymbol{e}_2$ (背叛) 之间选择. 第二阶段, 参与者必须在 $\boldsymbol{f}_1$ (惩罚背叛者) 与 $\boldsymbol{f}_2$ (不惩罚背叛者) 之间选择. 如 5.5 节, 假设每一个惩罚行为使得被惩罚者的收益减少 β, 惩罚者的收益减少 γ (其中 $\beta, \gamma>0$). 因此, 惩罚是一种昂贵的行为: 利己的参与者应该避免. 在这样的两阶段博弈中 (首先是贡献与否, 其次是惩罚与否), 有四种不同的策略. 亲社会策略 $\boldsymbol{G}_1=\boldsymbol{e}_1\boldsymbol{f}_1$ 贡献并惩罚. 矛盾策略 $\boldsymbol{G}_2=\boldsymbol{e}_2\boldsymbol{f}_1$ 背叛并惩罚所有背叛的对手. 利己策略 $\boldsymbol{G}_3=\boldsymbol{e}_2\boldsymbol{f}_2$ 既不贡献也不惩罚. 最后, 策略 $\boldsymbol{G}_4=\boldsymbol{e}_1\boldsymbol{f}_2$ 贡献, 但不惩罚背叛者. 使用这种策略的参与者可以看作是*二阶剥削者*, 搭了其他参与者惩罚的便车.

这一节只讨论 SR 情形. 每个参与者都得到 $N-1$ 个对手贡献的平均收益

$$B=\frac{rc}{N}(N-1)(x_1+x_4). \tag{6.4}$$

贡献与惩罚所产生的成本使得 G_i 类型参与者的平均净收益为 P_i, 其中

$$P_1 = B - c\left(1 - \frac{r}{N}\right) - (N-1)\gamma\,(x_2 + x_3)\,, \tag{6.5}$$

$$P_2 = B - (N-1)\beta\,(x_1 + x_2) - (N-1)\gamma\,(x_2 + x_3)\,, \tag{6.6}$$

$$P_3 = B - (N-1)\beta\,(x_1 + x_2)\,, \tag{6.7}$$

$$P_4 = B - c\left(1 - \frac{r}{N}\right). \tag{6.8}$$

考虑复制方程 $\dot{x}_i = x_i(P_i - \bar{P})$, 其中 $\bar{P} = \sum x_i P_i$ 是平均收益. 由于 $P_1 + P_3 = P_2 + P_4$, x_1x_3/x_2x_4 是不变运动. 因此, 集合

$$W_K = \{\boldsymbol{x} \in S_4 : x_1x_3 = Kx_2x_4\} \quad (K > 0) \tag{6.9}$$

为状态空间 S_4 中的一个层状的不变流形, 如图 5.1 所示. 这样, 研究这些二维流形上的动力学行为就足够了, 这些流形是由边 $\boldsymbol{G}_1 - \boldsymbol{G}_2 - \boldsymbol{G}_3 - \boldsymbol{G}_4 - \boldsymbol{G}_1$ 张成的鞍面.

如 5.5 节所述, 这些面的内部没有驻点. 边界 $\boldsymbol{G}_1\boldsymbol{G}_2$ 上的流指向 $\boldsymbol{G}_1$, 边界 $\boldsymbol{G}_2\boldsymbol{G}_3$ 和 $\boldsymbol{G}_4\boldsymbol{G}_3$ 上的流指向 $\boldsymbol{G}_3$. 边界 $\boldsymbol{G}_1\boldsymbol{G}_4$ 由驻点组成. 为使问题更为有趣, 我们往往假设

$$(N-1)\beta > c\left(1 - \frac{r}{N}\right). \tag{6.10}$$

此条件表明, 如果所有对手都惩罚, 则施加给一个背叛者的总惩罚高于其向公共利益贡献的净成本. 这样, 点

$$\mathbf{Q} := \left(\frac{(N-r)c}{\beta N(N-1)}, 0, 0, 1 - \frac{(N-r)c}{\beta N(N-1)}\right) \tag{6.11}$$

位于边界 $\boldsymbol{G}_1\boldsymbol{G}_4$ 上. 没有 $\boldsymbol{G}_2$ 和 $\boldsymbol{G}_3$[①]时, $\boldsymbol{G}_1\boldsymbol{Q}$ 上的增长率 $P_i - \bar{P}$ 为负. 因此, 这些点是饱和的, 即为纳什均衡. 而介于 $\boldsymbol{Q}$ 与 $\boldsymbol{G}_4$ 之间的点则不是. 这得到了如图 5.3 的相图. 这样, 对所有初值来说, 解或者流向 $\boldsymbol{G}_3$, 或者流向 $\boldsymbol{G}_1\boldsymbol{Q}$ 段. 如果随机冲击偶尔扰动系统, 则从长远来看, 利己策略 $\boldsymbol{G}_3$ 将胜出.

6.4 声 誉

假设有小概率 μ 使得 (i) 参与者偶尔知道对手的类型, 且 (ii) 如果背叛没有被惩罚的风险, 即如果所有的对手都是非惩罚的类型 $\boldsymbol{G}_3$ 或 $\boldsymbol{G}_4$, 则通常先选择贡献的参与者会改变其主意而决定不贡献. 这将改变收益值. 策略 $\boldsymbol{G}_i$ 的参与者现在的

①译注: 原文此处为 "$\boldsymbol{G}_4$".

期望收益 $P_i(\mu)$ 为

$$P_1(\mu) = B - c\left(1 - \frac{r}{N}\right)[1 - \mu\,(x_3 + x_4)^{N-1}] \\ - (N-1)\,\gamma\,(x_2 + x_3), \tag{6.12}$$

$$P_2(\mu) = B - (N-1)\,\beta\,(x_1 + x_2) - (N-1)\,\gamma\,(x_2 + x_3), \tag{6.13}$$

$$P_3(\mu) = B - (N-1)\frac{rc}{N}\mu(x_1 + x_4)(x_3 + x_4)^{N-2} \\ - (N-1)\,\beta\,(x_1 + x_2), \tag{6.14}$$

$$P_4(\mu) = B - (N-1)\frac{rc}{N}\mu(x_1 + x_4)(x_3 + x_4)^{N-2} \\ -c\left(1 - \frac{r}{N}\right)[1 - \mu\,(x_3 + x_4)^{N-1}], \tag{6.15}$$

其中 B 保持不变, 见等式 (6.4). 事实上, 非惩罚者的收益 P_3 和 P_4 由于贡献者改变主意而减小: 这会出现在 $N-1$ 个对手中有一个贡献者而其他 $N-2$ 个对手是非惩罚者的情形. 当所有 $N-1$ 个对手都是非惩罚者时, 贡献者的收益 P_1 和 P_4 将会改变.

又一次 $P_1(\mu) + P_3(\mu) = P_2(\mu) + P_4(\mu)$. 因此, W_K 是不变流形. 对于小 $\mu > 0$, 边界 $\boldsymbol{G}_1\boldsymbol{G}_2$, $\boldsymbol{G}_2\boldsymbol{G}_3$ 和 $\boldsymbol{G}_3\boldsymbol{G}_4$ 上流的方向不变, 但边界 $\boldsymbol{G}_1\boldsymbol{G}_4$ 不再由驻点组成: 现在这条边上的流指向 $\boldsymbol{G}_1$. 通过线性化可以看到, 在每一个 W_K 内, 顶点 $\boldsymbol{G}_1$ 和 $\boldsymbol{G}_3$ 是渊, 而 $\boldsymbol{G}_2$ 和 $\boldsymbol{G}_4$ 是源. 因此, 在每一个 W_K 的内部至少存在一个驻点. 而且, 这样的驻点只有一个 (于是这个点必须是鞍点, 见图 5.5). 事实上, 这个驻点必满足 $P_1(\mu) = P_2(\mu)$. 令 $y = x_3 + x_4$ (惩罚者的频度), 且

$$f(y) := P_1(\mu) - P_2(\mu), \tag{6.16}$$

我们有

$$f(y) = \mu\frac{(N-r)\,c}{N}y^{N-1} - \beta(N-1)y + \left[\beta(N-1) - \frac{(N-r)c}{N}\right]. \tag{6.17}$$

由于 f 严格凸, $f(1) < 0$ 且 $f(0) > 0$. 因此, 方程 $f(y) = 0$ 在 $]0,1[$ 有唯一解 $y = \hat{y}$. 另外, 驻点必须满足 $P_2(\mu) = P_3(\mu)$, 这样,

$$\gamma z = \frac{rc\mu}{N}(1 - z)\hat{y}^{N-2}, \tag{6.18}$$

其中 $z = x_2 + x_3$ 为背叛者的频度. 这就确定了 z. 注意, $\mu \to 0$ 时, W_K 上的内部驻点收敛到 $\boldsymbol{Q}$.

因此, 在状态空间 S_4 出现了双稳态: 除了一个零测集, 所有解或者趋于亲社会状态或者趋于利己状态. 数值模拟表明, 即使对非常小的 μ, 亲社会策略 $\boldsymbol{G}_1$ 的吸引盆总是存在的.

这个模型和其他几个模型都没有解释社会惩罚的出现. 如果所有参与者都愿意惩罚, 即处于亲社会状态 G_1, 则背叛者显然没有机会侵入. 但在利己状态 G_3, 亲社会参与者不能立足. 一个下决心惩罚所有背叛者的参与者必定为了惩罚疲于奔命. 这种行为代价太大而不太可能被模仿.

6.5 有限群体

为了从另外的角度考察惩罚系统, 考虑规模为 M 的有限群体. 随机选取有 N 个参与者的样本进行公共利益博弈. 只考虑三种策略, 用 X, Y 和 Z 表示. X 参与者总是贡献, 但不惩罚; Y 参与者既不贡献也不惩罚; 而 Z 参与者贡献并惩罚. 在前面的模型中, 对应于 G_4 (二阶剥削者), G_3 (既不贡献也不惩罚的利己参与者) 和 G_1 (既贡献又惩罚背叛者的亲社会参与者). 为了简单起见, 不考虑背叛但又惩罚背叛者的矛盾策略 G_2.

假设一个特别简单的模仿机制. 随机选取两个参与者并比较其收益. 低收益的参与者采用另一个参与者的策略. 如果两个参与者的收益相同, 则掷硬币决定谁模仿谁. 如果这些更新不够充分, 则收益值 (依赖于群体中进行公共利益博弈的群体的随机样本) 非常接近期望值.

对应的随机过程是一个马尔可夫链, 同质状态 $(M,0,0)$, $(0,M,0)$ 和 $(0,0,M)$ 为其吸收状态, 分别记为 *AllX*, *AllY* 和 *AllZ*. 模仿不能导致这些状态改变. 另外, 假设参与者可以偶尔不通过模仿而随机选择另一种策略. 这个 "突变" 将同质状态变为有一个不同策略者的状态. 下一步, 又进入模仿过程. 此时, 或者这个不同策略者切换回先前的同质策略, 或者群体中的其他参与者最终选择新策略. 这样, 群体依然保持同质直到下一次随机突变发生, 等等.

假设这些随机突变不经常发生, 我们可以有效地进行时间尺度分离: 模仿比突变快得多. 这种 "不可转移的情形" 已经在 2.17 节描述过. 从一种同质状态到另一种同质状态的转移可以由具有三种状态 *AllX*, *AllY* 和 *AllZ* 的马尔可夫链描述, 由

$$\begin{pmatrix} \frac{1}{2}-\frac{1}{2M} & \frac{1}{2} & \frac{1}{2M} \\ 0 & 1 & 0 \\ \frac{1}{2M} & 0 & 1-\frac{1}{2M} \end{pmatrix} \tag{6.19}$$

给出. 例如, 群体在状态 *AllX*, 一个参与者随机选择另一种策略. 该策略有 1/2 的概率为背叛者的策略 Y. 由于 Y 参与者的收益总是比 X 参与者的收益好, 群体将会以状态 *AllY* 结束. 以相同的概率 1/2, 随机突变也可以引入亲社会策略 Z. 在只有 X 参与者和 Z 参与者的群体中, 所有参与者收益一样好. 最终, 所有个体采取 Z 策略 (通过中性漂变) 的概率是 $1/M$. *AllX* 到 *AllX* 的转移概率使得矩阵的行之和为

1. 在第二行中: 对于一个 *AllY* 的群体, 单个 X 个体会被剥削, 没有常驻者收益好. 因此, 它不会被模仿, 下一次又会回到 Y 策略. 类似地, 单个 Z 参与者不会比常驻者收益好, 在 SR 情形有

$$c\left(1-\frac{r}{N}\right)+\gamma(N-1)>\frac{N-1}{M-1}\left(\beta-\frac{rc}{N}\right), \tag{6.20}$$

在**OO**情形有

$$c+\gamma(N-1)>\frac{N-1}{M-1}\left(\beta-\frac{cr}{N-1}\right). \tag{6.21}$$

如果群体规模 M 足够大, 则这两个条件都能自然满足. 最后, 考虑个体从 *AllZ* 状态以相等的概率突变到策略 X 或策略 Y. 前一种情形, 又会出现中性漂变, 其固定概率为 $1/M$. 后一种情形, 固定概率为 0, 这只要假设单个背叛者收益不会比常驻的惩罚者好. 这在 SR 的情形为

$$(N-1)\beta>\frac{N-1}{M-1}\gamma+c\left(1-\frac{r}{N}\right)\left(\frac{M-N}{M-1}\right), \tag{6.22}$$

在**OO**情形为

$$c\left(1+\frac{r}{M-1}\right)<(N-1)\left(\beta-\frac{\gamma}{M-1}\right). \tag{6.23}$$

当 $M\to\infty$ 时, 不等式 (6.22) 简化为 $(N-1)\beta>(1-\frac{r}{N})c$ (即 (6.10)), (6.23) 简化为 $(N-1)\beta>c$.

容易看出, 由矩阵 (6.19) 给出的马尔可夫链有唯一的平稳分布, 即 $(0,1,0)$. 这表明, *AllY*, 即纯背叛者状态, 是不可避免的结果.

6.6 志愿者出列

现在, 考虑一种不参加公共利益博弈的策略, 用 W 表示. 采取这类策略的参与者将获得不依赖于其他参与者的收益 σ. 假设

$$0<\sigma<(r-1)c. \tag{6.24}$$

这意味着, 这类自给自足不参加共同努力的参与者, 其收益比在一个所有成员都贡献的公共利益博弈中的参与者所获得的低, 但比在一个只有背叛者参与的公共利益博弈中的参与者所获得的高. 进一步假设单个参与者不能只靠自己完成公共利益博弈. 需要至少另一个参与者的参加.

描述四种同质状态 *AllX*, *AllY*, *AllZ* 和 *AllW* 之间转移的马尔可夫链为

$$\begin{pmatrix} \frac{2}{3}-\frac{1}{3M} & \frac{1}{3} & \frac{1}{3M} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{1}{3M} & 0 & 1-\frac{1}{3M} & 0 \\ \frac{1}{6} & 0 & \frac{1}{6} & \frac{2}{3} \end{pmatrix}. \tag{6.25}$$

只需要看最后一行和最后一列. (其他项可以如矩阵 (6.19) 中的一样得到, 现在随机突变到三种选择之一的概率为 1/3.) 显然, 单个 W 参与者的收益不会好于 X 常驻者和 Z 常驻者, 但比 Y 常驻者好. 在一个全为 W 参与者的群体中, 单个参与者转移到其他策略的收益与常驻者一样, 因为这个参与者不能参加任何公共利益博弈. 因此, 通过比较收益, 异议者与常驻者相互模仿的概率是一样的. 第二个常驻者有 1/2 的概率采取异议者的策略. 如果异议者的策略为 Y 策略, 则其收益不会好于 W 常驻者, 而如果是 X 策略或 Z 策略, 则比常驻者好.

容易看到, 马尔可夫链 (6.25) 有唯一的平稳分布

$$(p, p, 1-3p, p), \tag{6.26}$$

其中

$$p = \frac{2}{M+8}. \tag{6.27}$$

对 $M = 100$, 这意味着几乎有 95% 的时间, 群体只由亲社会类型组成. 背叛者下降到 2% 以下, 而在前一节它们主导着强制性公共利益博弈.

6.7 选择弃权

为了更好地理解自愿参加的作用, 考虑不能惩罚的情形. 此时, 只有策略 X, Y 和 W, 而没有策略 Z. 转移矩阵为

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}, \tag{6.28}$$

对应的平稳分布为

$$\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right). \tag{6.29}$$

这意味着, 从长远来看, 有一半的时间所有的参与者都参加到全都贡献或是全都背叛的博弈中, 有一半的时间没有参与者参加. 矩阵 (6.28) 呈现了一个石头 — 剪刀 — 布的结构: *AllX* 只能向 *AllY* 突变, *AllY* 只能向 *AllW* 突变, 而 *AllW* 只能向 *AllX* 突变. 正是这个循环避免了一个全是背叛者状态的死结.

现在, 通过考虑无限大群体的复制系统来分析这个循环, 群体中包括合作者 (贡献但不惩罚), 背叛者和自给自足参与者. 这里不考虑惩罚者. 用 x, y 和 w 表示三种策略的频度, 其期望收益值分别记为 P_x, P_y 和 P_w. 显然,

$$P_w = \sigma. \tag{6.30}$$

一个参与者与 h 个对手进行公共利益博弈的概率为

$$\binom{N-1}{h}(1-w)^h w^{N-1-h}, \quad h = 0, \cdots, N-1, \tag{6.31}$$

其中, m 个为合作者的概率为

$$\binom{h}{m}\left(\frac{x}{1-w}\right)^m \left(\frac{y}{1-w}\right)^{h-m}, \tag{6.32}$$

其中 $m = 0, \cdots, h$ $(h > 0)$. 如果 $h = 0$, 则此参与者被迫选择非参加策略, 其收益为 σ.

首先考虑 OO 情形, 从 h 个共同参与者中得到的期望收益为

$$\sum_{m=0}^{h} \frac{rcm}{h}\binom{h}{m}\left(\frac{x}{1-w}\right)^m \left(\frac{y}{1-w}\right)^{h-m} = \frac{rcx}{1-w}, \tag{6.33}$$

它独立于 h $(h = 1, \cdots, N-1)$. 因此, 一个背叛者从公共利益博弈中获得的收益为

$$P_y = \sigma w^{N-1} + \frac{rcx\left(1-w^{N-1}\right)}{1-w}. \tag{6.34}$$

此式再减去 $c(1-w^{N-1})$ 为合作者的收益. 因此

$$P_y - P_x = c\left(1-w^{N-1}\right), \tag{6.35}$$

此式为正.

在 SR 的情形, 当捐赠的一部分返回给贡献者时, 表达式会稍微复杂一些. 背叛者在有 h 个共同参与者的群体中获得的收益为 $rcm/(h+1)$, 其中 m 为合作者的数量. 因此, 背叛者的期望收益为

$$\left(\frac{rch}{h+1}\right)\left(\frac{x}{1-w}\right). \tag{6.36}$$

这样,

$$P_y = \sigma w^{N-1} + rc\frac{x}{1-w}\sum_{h=0}^{N-1}\binom{N-1}{h}(1-w)^h w^{N-h-1}\left(\frac{h}{h+1}\right) \tag{6.37}$$

($h=0$ 的项不影响和式). 由等式

$$\binom{N-1}{h} = \binom{N}{h+1}\frac{h+1}{N} \tag{6.38}$$

可得

$$P_y = \sigma\, w^{N-1} + rc\frac{x}{1-w}\left(1-\frac{1-w^N}{N(1-w)}\right). \tag{6.39}$$

在一个有 h 个对手参加的 SR 公共利益博弈中, 从合作变为背叛所产生的成本为 $c(1-r/(h+1))$. 因此

$$P_y - P_x = \sum_{h=1}^{N-1} c\left(1-\frac{r}{h+1}\right)\binom{N-1}{h}(1-w)^h w^{N-h-1}. \tag{6.40}$$

如前, 利用等式 (6.38) 可得

$$P_y - P_x = cF_N(w), \tag{6.41}$$

其中

$$F_N(w) := 1 + (r-1)w^{N-1} - \left(\frac{r}{N}\right)\left(\frac{1-w^N}{1-w}\right). \tag{6.42}$$

背叛者相对合作者的优势只依赖于自给自足参与者的数量 w.

$P_y - P_x$ 的符号决定了是否值得从合作变为背叛, $F_N(w)=0$ 是平衡条件. 当 $r \leqslant 2$ 时 F_N 没有实根, 当 $r>2$ 时, 在区间 $[0,1]$ 只有一个实根 $\hat{w}$. 为此, 考虑函数 $G(w)=F_N(w)(1-w)$, 在区间 $[0,1]$ 内它与 F_N 有相同的实根, 注意到 $G(0)=1-r/N>0$ 而 $G(1)=0$. 当 $r>2$ 时, 函数 G 在 $w=1$ 有一个局部最大值. 事实上, $G'(1)=0$, 且 $G''(1)=(2-r)(N-1)$. 此外, $G''(w)=w^{N-3}(N-1)[(N-2)(r-1)-w(Nr-N-r)]$ 在 $[0,1]$ 最多改变一次符号. 因此, 当 $r>2$ 时, 存在一个自给自足参与者频度的临界值 $\hat{w}$, 大于这个值时合作者的收益比背叛者好, 如图 6.1 所示.

群体的平均收益为 $\bar{P} = xP_x + yP_y + wP_w$, 利用 $y=1-x-w$ 可以将其改写为

$$\bar{P} = x(P_x - P_y) + w(\sigma - P_y) + P_y = -x(P_y - P_x) + (1-w)(P_y - \sigma) + \sigma. \tag{6.43}$$

这样, 对 SR 和 OO 两种情况都有

$$\bar{P} = \sigma + [(r-1)xc - (1-w)\sigma](1-w^{N-1}). \tag{6.44}$$

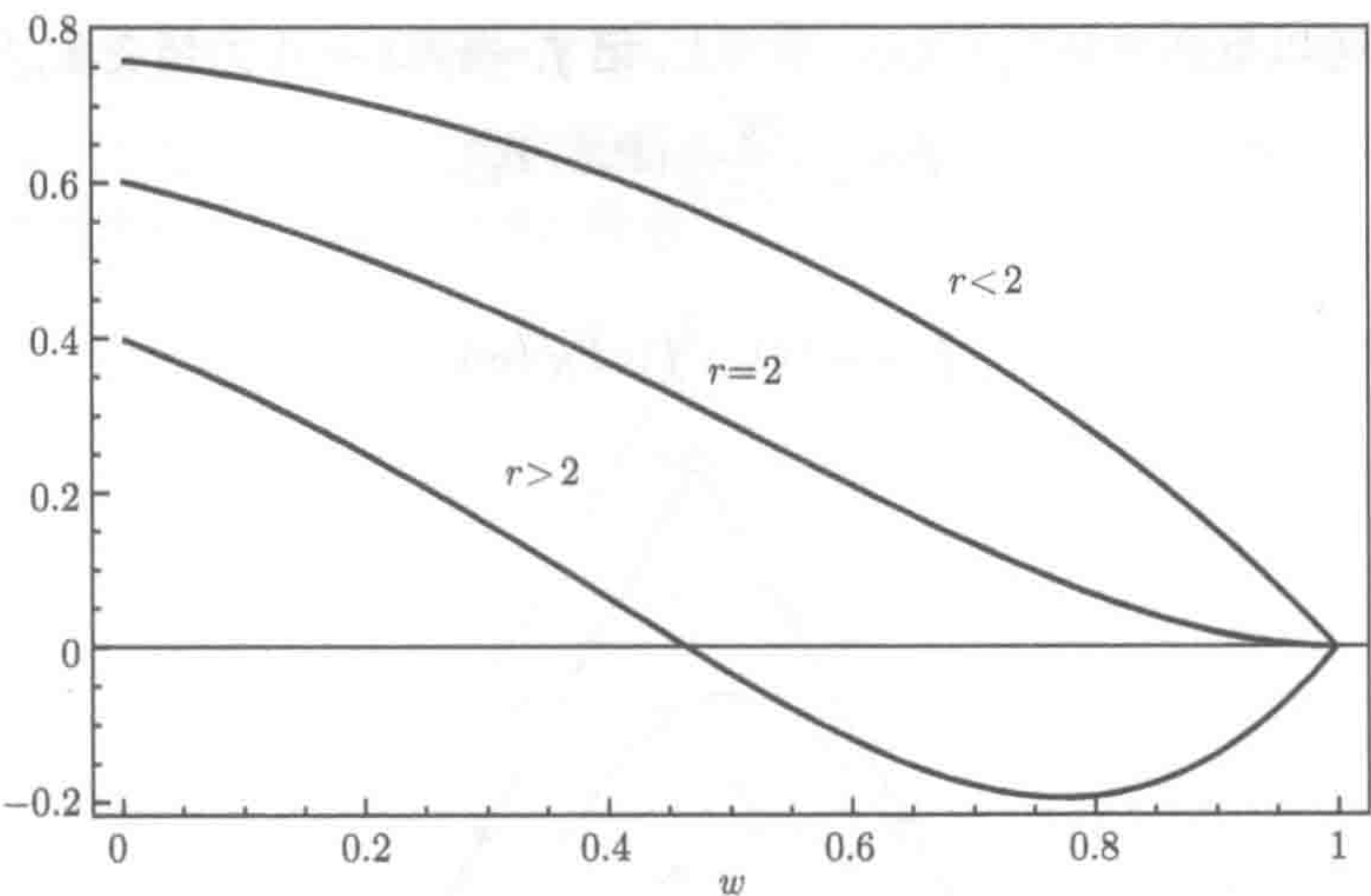


图 6.1 函数 $w \mapsto (1-w)F_N(w)$ 的图像, 其中 F_N 由 (6.42) 给出, $N=5$

6.8 石头—剪刀—布动力学

现在分析三策略 X, Y, W, 即合作者、背叛者与自给自足参与者的复制动力学. 在 S_3 的边界上, SR 和 OO 的动力学行为是相同的. 单形 $S_3 = \{(x,y,w): x,y,w \geqslant 0, x+y+w=1\}$ 的顶点, 即同质状态 *AllX*, *AllY* 以及 *AllW* 都是驻点. S_3 的边界上没有驻点. 事实上, 轨道从 *AllX* (只有合作者) 指向 *AllY* (只有背叛者), 从 *AllY* 指向 *AllW* (自给自足参与者), 从 *AllW* 指向 *AllX*. 因此, S_3 的边界是一个异宿环.

在 OO 情形, 总有 $P_x < P_y$. 状态空间 S_3 的内部没有驻点: 当 $t \to \pm\infty$ 时, 所有轨道趋于 *AllW* (图 6.2).

这一节余下的部分, 我们只考虑 SR 情形. 当 $r \leqslant 2$ 时, (6.42) 定义的函数 F_N 在 $]0,1[$ 内没有实根. 从而, 在 S_3 的内部有 $P_x < P_y$: 这得到了与 OO 情形相同的动力学行为, 如图 6.2 所示. 因此, 现在假设 $r > 2$.

此时, F_N 有一个实根 $\hat{w} \in [0,1]$. 利用 $P_y = P_w = \sigma$, 在 S_3 的内部可得唯一驻点 $\boldsymbol{Q} = (\hat{x}, \hat{y}, \hat{w})$, 其中

$$\hat{x} = \frac{\sigma}{c(r-1)}(1-\hat{w}), \tag{6.45}$$

以及

$$\hat{y} = \left(1 - \frac{\sigma}{c(r-1)}\right)(1-\hat{w}). \tag{6.46}$$

为了分析 S_3 内部的动力学行为, 将复制方程改写为汉密尔顿系统的形式是有

效的, 这样可以得到一个不变运动. 事实上, 记 $f = x/(x+y)$ 为新变量, 则可得

$$\dot{f} = \frac{xy}{(x+y)^2}(P_x - P_y). \tag{6.47}$$

即有

$$\dot{f} = -f(1-f)\,cF_N(w). \tag{6.48}$$

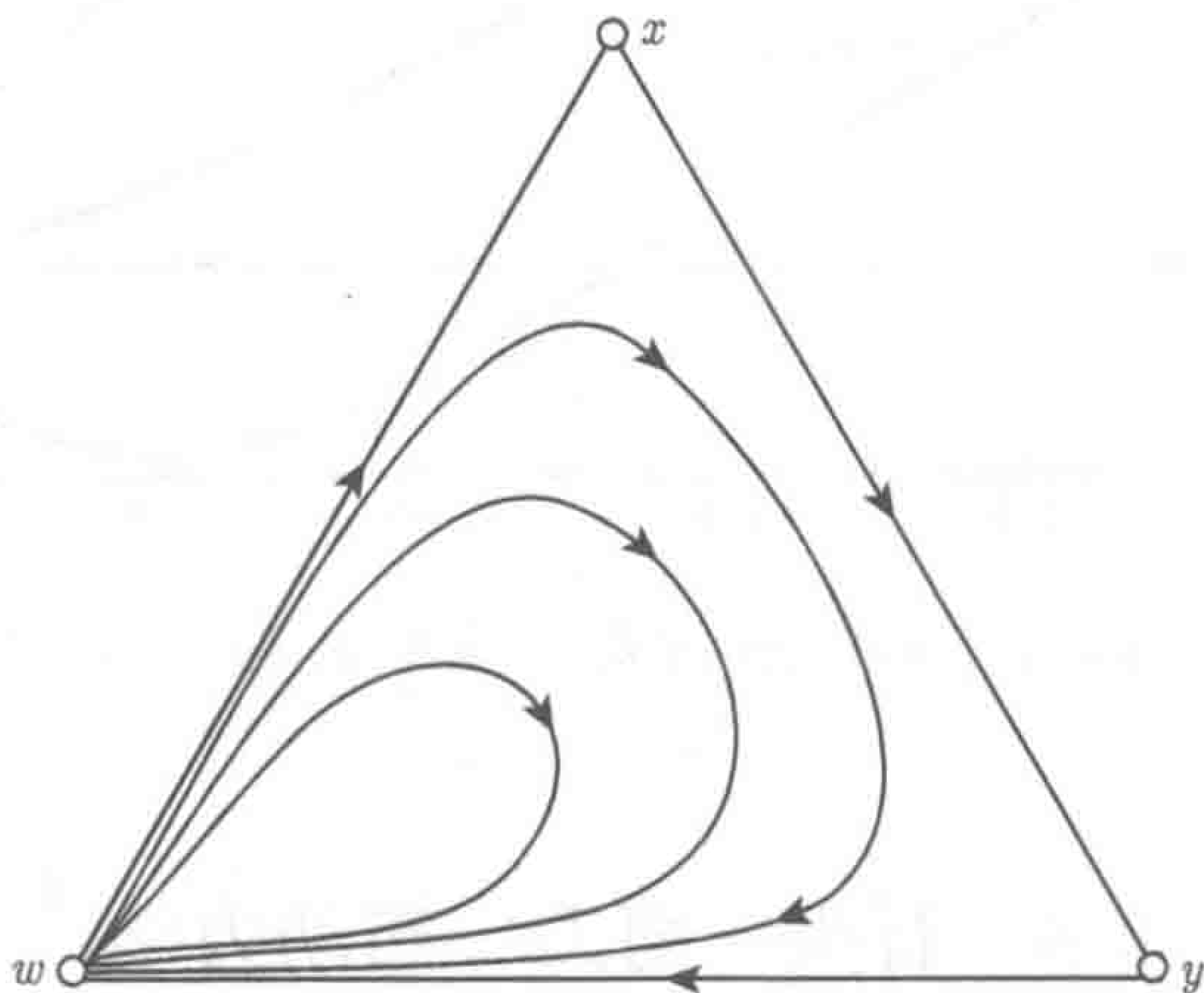


图 6.2　OO 情形或 SR 情形 ($1 < r \leqslant 2$) 的复制动力学. 在 S_3 的内部 $t \to \pm\infty$ 时所有轨道趋于 $w = 1$

利用等式 (6.44) 可得

$$\dot{w} = [\sigma - cf(r-1)]w(1-w)(1-w^{N-1}). \tag{6.49}$$

在等式 (6.49)[①] 和 (6.48) 的右端除以函数 $f(1-f)w(1-w)(1-w^{N-1})$, 此函数对所有 0 与 1 之间的 f 和 w 值为正, 它对应速度的改变, 但不影响轨道. 这样, 有

$$\dot{f} = \frac{-cF_N(w)}{w(1-w)(1-w^{N-1})} =: -g(w), \tag{6.50}$$

$$\dot{w} = \frac{\sigma - cf(r-1)}{f(1-f)} =: l(f). \tag{6.51}$$

引入 $H := G + L$, 其中 $G(w)$ 和 $L(f)$ 是 $g(w)$ 和 $l(f)$ 的原函数, 则可得汉密尔顿系统

$$\dot{f} = -\frac{\partial H}{\partial w}, \tag{6.52}$$

$$\dot{w} = \frac{\partial H}{\partial f}. \tag{6.53}$$

①译注: 原文此处为 "(6.47)".

这个系统是保守的, 汉密尔顿量 H 在 $(\frac{\sigma}{c(r-1)}, \hat{w})$ 达到严格全局最大值. 因此, $(x, y, w) \in S_3$ 到 $(f, w) \in [0,1]$ 的变量变换表明, 复制系统内部唯一的平衡点 $\boldsymbol{Q}$ 是一个由闭轨围绕的稳定点, 如图 6.3 所示.

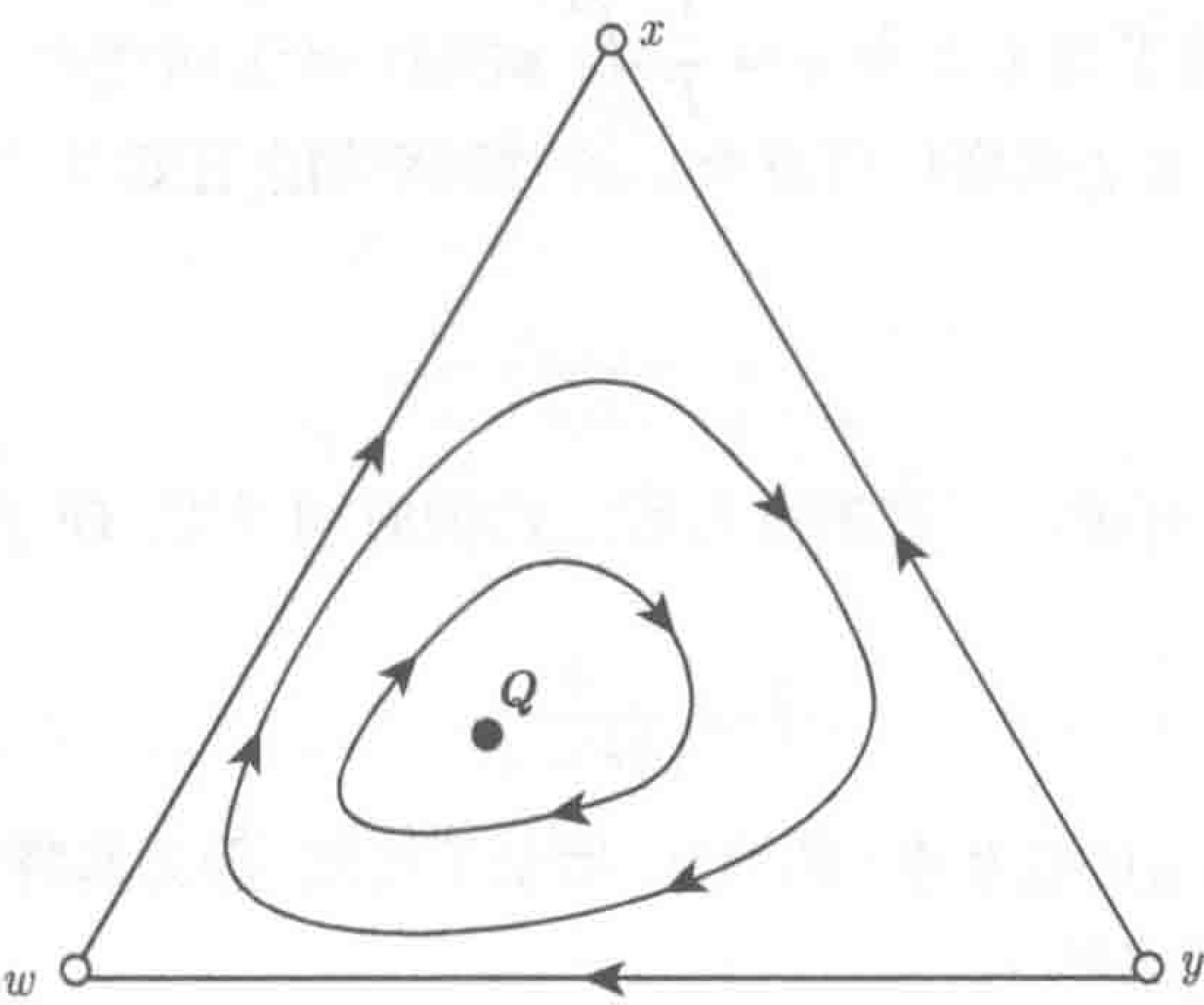


图 6.3 当 $r > 2$ 时, SR 情形的复制动力学. 驻点 $\boldsymbol{Q}$ 被周期轨包围

通过改变三个参数 N, r 和 σ, 可以使得 $\boldsymbol{Q}$ 处于单形内部任意一点 (图 6.4). 驻点 $\boldsymbol{Q}$ 在独立于样本容量 N 的直线

$$x = \frac{\sigma}{c(r-1)-\sigma} y \tag{6.54}$$

上. 当 N 增加时, $\boldsymbol{Q}$ 朝顶点 $AllW$ 移动. 极限 $N \to \infty$ 得到从 $AllW$ 流出并流入 $AllW$ 的同宿轨.

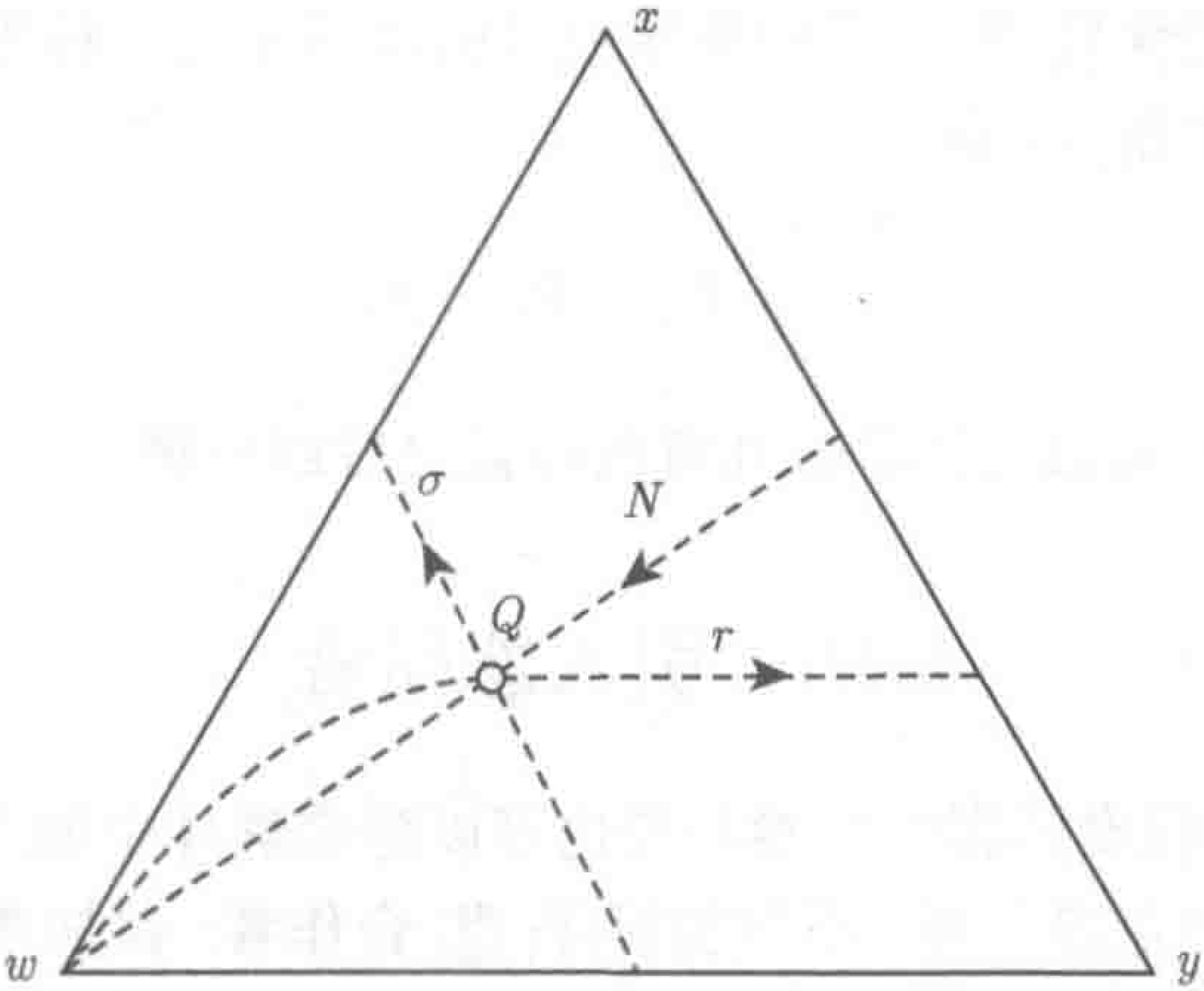


图 6.4 $\boldsymbol{Q}$ 对参数 N, σ 和 r 的依赖

6.9 时间平均

考虑自愿参加的、SR 情形公共利益博弈的复制动力学. 将函数 $v: S_3 \to R$ 沿 T 周期轨道上的时间平均定义为 $\bar{v} = \frac{1}{T}\int_0^T v(x(t), y(t), w(t))\mathrm{d}t$. 通常, 它依赖于初值条件, 但合作者与参加博弈的所有个体的时间平均之比等于平衡点 Q 处对应分量的比值:

$$\frac{\bar{x}}{\bar{x}+\bar{y}} = \frac{\sigma}{c(r-1)}. \tag{6.55}$$

进一步, 合作者与所有参加博弈的参与者之比的时间平均, 即 $\bar{f}$, 对应于时间平均之比:

$$\bar{f} = \frac{\sigma}{c(r-1)}. \tag{6.56}$$

有趣的是, r (公共利益的盈利率) 的增加, 助长了背叛, 因为其降低了实际参加公共利益博弈中合作者的比率 f,

为了得到等式 (6.55), 利用关系 $x = f(1-w)$. 在等式 (6.49) 两端除以 $w(1-w^{N-1})$, 可得

$$\int_0^T [\sigma(1-w) - c(r-1)x]\mathrm{d}t = \int_0^T \frac{\dot{w}\mathrm{d}t}{w(1-w^{N-1})} = p(w)\Big|_{w(0)}^{w(T)}. \tag{6.57}$$

其中, $p(w)$ 是 $[w(1-w^{N-1})]^{-1}$ 的原函数. 由于轨道都是闭的, 所以最后一项为零, 即可得 $\bar{x}$ 与 $1-\bar{w} = \bar{x}+\bar{y}$ 之间的比. 在等式 (6.49) 两端除以 $w(1-w)(1-w^{N-1})$, 则同样可得等式 (6.56).

由于复制方程的性质, 三个不同策略收益的时间平均是相等的, 并可以简化为自给自足参与者的收益, 即 σ:

$$\bar{P}_x = \bar{P}_y = \bar{P}_w = \sigma. \tag{6.58}$$

因此, 从长远看来, 没有谁会比非参加者的收益更好或更糟.

6.10 引入惩罚者

如 6.3 节, 现在假设博弈中的参与者也可以惩罚群体中的背叛者. 这样, 有四类参与者: ① 自给自足参与者: 不参加博弈; ② 合作者: 参加博弈并捐赠, 但不惩罚; ③ 背叛者: 参加博弈, 既不捐赠也不惩罚; 以及④ 惩罚者: 参加博弈, 捐赠并惩罚群体中的背叛者. 在无限群体中, 合作者、背叛者、惩罚者和自给自足参与者的

频度用 x, y, z 以及 w 表示. 在给定的大小为 N 的随机样本中, 它们的频数分别用 N_x, N_y, N_z 以及 N_w 表示 (其中 $N_x+N_y+N_z+N_w=N$ 而 $S=N_x+N_y+N_z$ 是参加公共利益博弈的参与者的数量). 与 6.4 节相比, 这里没有考虑声誉的作用.

到现在为止, 假设了在每一回合公共利益博弈之后, 惩罚者惩罚背叛者, 每一次惩罚行为使得惩罚者损失 γ, 被惩罚者损失 β, 所以惩罚者必须支付 γN_y, 而背叛者支付 βN_z. 这就是所谓的 "同级惩罚" (peer-punishment) 或 "私人惩罚" (private punishment), 记为 PP. 另外, 考虑所谓的 "惩罚机构" (sanctioning institution), 记为 SI. 此时, N_z 个惩罚者中的每个个体在公共利益博弈实际进行之前都支付 γ; 然后, 背叛者群体将被课以大小为 βN_z 的罚款. (如果没有背叛者, 则没有人会被惩罚, 惩罚者支付的费用不会返还; 否则, 背叛者平均分摊罚金, 因此每个背叛者支付 $\beta N_z/N_y$.)

总收益由公共利益项和惩罚项 (对自给自足参与者和合作者为 0) 构成. 容易看到, 在**PP**情形, 背叛者惩罚项的期望为

$$-\beta z(N-1), \tag{6.59}$$

惩罚者的惩罚项为

$$-\gamma y(N-1). \tag{6.60}$$

在 SI 情形, 背叛者的对应项为

$$-\beta\frac{z}{y}[1-(1-y)^{N-1}], \tag{6.61}$$

惩罚者的对应项为

$$-\gamma(1-w^{N-1}). \tag{6.62}$$

后者很显然: 除非样本中没有其他参与者想参加博弈, 否则, 一个惩罚者支付的费用为 γ. 现在考虑一个背叛者, 如果样本中另有 m 个背叛者, 则还剩 $N-m-1$ 个其他参与者, 其中 惩罚者的比率为 $z/(1-y)$. 这样, 恰好有 k 个是惩罚者的概率为

$$\binom{N-m-1}{k}\left(\frac{z}{1-y}\right)^k\left(1-\frac{z}{1-y}\right)^{N-m-1-k}, \tag{6.63}$$

其中惩罚为 $\boldsymbol{k}\beta/(m+1)$. 如果另有 m 个背叛者, 就得到一个条件期望

$$-\frac{\beta}{m+1}\left(\frac{z}{1-y}\right)(N-m-1). \tag{6.64}$$

这样, 惩罚项为

$$-\sum_{m=0}^{N-1}\left(\frac{\beta z}{1-y}\right)\frac{N-m-1}{m+1}\binom{N-1}{m}y^m(1-y)^{N-m-1}. \tag{6.65}$$

由于和式的最后一项为 0, 且有

$$\binom{N-1}{m}\frac{N-m-1}{m+1}=\binom{N-1}{m+1}, \tag{6.66}$$

因此, 惩罚项为

$$-\frac{\beta z}{y}\sum_{m=0}^{N-2}\binom{N-1}{m+1}y^{m+1}(1-y)^{N-1-(m+1)}, \tag{6.67}$$

即可得 (6.61).

6.11 同级惩罚的复制动力学

收益值是公共利益项和惩罚项的和. 惩罚项已经在前一节得到.

在 SR 情形, 自给自足参与者从公共利益博弈中得到收益 σ, 而背叛者得到的收益为

$$\sigma w^{N-1}+rc(x+z)H_N(w). \tag{6.68}$$

参见等式 (6.39), 合作者和惩罚者都得到收益

$$\sigma w^{N-1}+c(r-1)(1-w^{N-1})-rcyH_N(w), \tag{6.69}$$

参见等式 (6.41), 其中

$$H_N(w)=\frac{1}{1-w}\left(1-\frac{1-w^N}{N(1-w)}\right). \tag{6.70}$$

在 OO 情形, 背叛者从公共利益博弈中获得的收益为

$$\sigma w^{N-1}+\frac{rc(x+z)(1-w^{N-1})}{1-w}. \tag{6.71}$$

合作者和惩罚者从公共利益博弈中获得的收益相同, 但比背叛者少 $c(1-w^{N-1})$.

现在考虑 OO-PP 情形的复制动力学. 从所有收益中去掉公共项 σw^{N-1}, 可以分别得到自给自足参与者、背叛者、合作者和惩罚者的期望收益

$$P_w=(1-w^{N-1})\sigma, \tag{6.72}$$

$$P_y=(1-w^{N-1})\left(rc\frac{x+z}{1-w}\right)-\beta z(N-1), \tag{6.73}$$

$$P_x=(1-w^{N-1})\left(rc\frac{x+z}{1-w}-c\right), \tag{6.74}$$

以及

$$P_z=(1-w^{N-1})\left(rc\frac{x+z}{1-w}-c\right)-\gamma y(N-1). \tag{6.75}$$

因为 $P_z < P_x$, 所以在单形 S_4 的内部没有驻点. 所有轨道趋于边界. 在面 $z=0$ 上, 存在石头 — 剪刀 — 布博弈: 自给自足参与者被合作者占优, 合作者被背叛者占优, 背叛者被自给自足参与者占优. 在这个面的内部, 所有轨道都是同宿轨, 如果时间趋于 $\pm\infty$, 则轨道趋于自给自足参与者状态 (图 6.2).

容易看到, 如果 $c < \beta(N-1)$, 则背叛者相对自给自足参与者占优, 惩罚者和背叛者形成双稳系统. 合作者与惩罚者的边界 $(y=w=0)$ 由驻点组成, 满足

$$z > \frac{c}{\beta(N-1)} \tag{6.76}$$

的部分是饱和的, 因此是纳什均衡.

现在考虑 SR-PP 的情形. 通过计算可得 $P_w = \sigma$,

$$P_y = \sigma w^{N-1} + rc(x+z)H_N(w) - \beta z(N-1), \tag{6.77}$$

$$P_x = \sigma w^{N-1} + (r-1)c(1-w^{N-1}) - rcyH_N(w), \tag{6.78}$$

以及

$$P_z = \sigma w^{N-1} + (r-1)c(1-w^{N-1}) - rcyH_N(w) - \gamma y(N-1). \tag{6.79}$$

边界 $y=w=0$ 由驻点组成, 所有满足

$$z > \frac{c}{\beta(N-1)}\left(1-\frac{r}{N}\right) \tag{6.80}$$

的部分为纳什均衡. 在面 $z=0$ 上, 边界为一个异宿环. 当 $r \leqslant 2$ 时, 其动力学行为与 OO-PP 的情形相同. 当 $r > 2$ 时, 面 $z=0$① 包含一个由周期轨围绕的驻点 Q, 如图 6.3 所示. Q 点是饱和的, 因此是纳什均衡. 事实上, 在 Q 点有 $P_x = (r-1)c(1-w^{N-1}) + \sigma w^{N-1} - rcyH_N(w) = \bar{P} = P_w = \sigma$, 这样, $P_z - \bar{P} = P_z - P_x = -\gamma(N-1)y < 0$. 此外, 在面 $z=0$ 上, 任何以 T 为周期的闭轨 o 吸引 S_4 内部与之相邻的轨道, 即, "横截增长率" $P_z - \bar{P}$ 的时间平均为负. 这可以如前得到, 沿着轨道 o 的时间平均满足等式 $\hat{P}_x = \hat{P}_y = \hat{P}_w = \hat{\bar{P}} = \sigma$, 因此

$$\hat{P}_z - \hat{\bar{P}} = \hat{P}_z - \hat{P}_x = -\gamma(N-1)\frac{1}{T}\int_0^T y\mathrm{d}t < 0. \tag{6.81}$$

所以, 在这个意义上来说, 周期轨 o 是饱和的, 即横截稳定, 甚至是吸引的. 注意到, 对非常大的轨道, 其在靠近 $AllW$ 时花了大部分时间. 在这一点的横截特征值为 0.

由于在状态空间内部有 $P_z < P_x$, 所以, 所有轨道趋于边界. 边界上有两个集合吸引内部轨道: 一个集合是合作者 – 惩罚者的边界上满足不等式 (6.76) 的部分,

①译注: 原文此处为 "$w=0$".

另一个集合是充满周期轨的面 $z=0$. 由于退化的动力学行为 (具有驻点和周期轨的连续统), 复制系统是结构不稳定的: 非常小的扰动能产生截然不同的结果. 因此, 对于预测而言没有价值. 此时, 考虑有限群体的随机过程更为可取.

6.12 有限群体

考虑规模为 M 的有限群体. 用 X, Y, Z 以及 W 表示合作者, 背叛者, 惩罚者和自给自足参与者的数量 $(X+Y+Z+W=M)$. 容量为 N 的随机样本中有 N_x 个合作者, N_y 个背叛者, N_z 个惩罚者和 N_w 个自给自足参与者的概率服从超几何分布

$$\frac{\binom{X}{N_x}\binom{Y}{N_y}\binom{Z}{N_z}\binom{W}{N_w}}{\binom{M}{N}}. \tag{6.82}$$

由于考虑有限群体, 因此必须假设不放回取样.

虽然有些繁杂, 但这些类型期望收益的计算是直接的. 例如, 在 SR-PP 情形的期望收益为 $P_w=\sigma$,

$$P_y=\frac{W_{N-1}}{(M-1)_{N-1}}\sigma+H-\frac{Z}{M-1}(N-1)\beta, \tag{6.83}$$

$$P_x=\frac{W_{N-1}}{(M-1)_{N-1}}\sigma+H-F(W)c, \tag{6.84}$$

$$P_z=\frac{W_{N-1}}{(M-1)_{N-1}}\sigma+H-F(W)c-\frac{Y}{M-1}(N-1)\gamma, \tag{6.85}$$

其中,

$$W_N=W(W-1)\cdots(W-N+1), \tag{6.86}$$

$$H=rc\frac{X+Z}{M-W-1}\left[1-\frac{1}{N(M-W)}\left(M-(W-N-1)\frac{W_{N-1}}{(M-1)_{N-1}}\right)\right], \tag{6.87}$$

以及

$$F(W)=1-\frac{r}{N}\frac{M-N}{M-W-1}+\frac{W_{N-1}}{(M-1)_{N-1}}\left(\frac{r}{N}\frac{W+1}{M-W-1}+r\frac{M-W-2}{M-W-1}-1\right). \tag{6.88}$$

据此, 可以计算每个学习过程的转移概率: 由于群体中包含四种策略类型, 就产生了一个有 $\binom{M+3}{3}$ 种 (X,Y,Z,W) 状态的马尔可夫链. 对合理的 M 值, 如 $M=100$ 或 $M=1000$, 状态种类相当多. 如果没有突变, 群体将以同质状态结束, 所有个体或者成为合作者 $(X=M)$ 或者成为背叛者 $(Y=M)$ 等等. 如 2.17 节所述, 如果在模仿过程中引入一个非常小的 "突变项", 则这些状态不再是吸收的, 我们能分离时间尺度. 这种 "不可转移" 的方法, 将过程简化为一个马尔可夫链, 其描述了四种同

质状态 *AllX*, *AllY*, *AllZ* 和 *AllW* 之间的转移. 对应的左特征向量 $(\pi_x, \pi_y, \pi_z, \pi_w)$ 产生了四种同质状态的平稳分布.

6.4 节已经用过这种方法, 即考虑一个非常简单的模仿过程. 本节, 通过将 Moran 过程作为学习规则 (2.15 至 2.17 节) 来测试结果的稳健性 (robustness).

由等式 (2.78), 固定概率可以通过形如

$$\rho_{xy} = \frac{1}{1 + \sum_{k=1}^{M-1} \prod_{X=1}^{k} \frac{1 - s + sP_{XY}}{1 - s + sP_{YX}}} \tag{6.89}$$

的公式得到, 其中 P_{XY} 表示合作者在有 X 个合作者和 $Y = M - X$ 个背叛者的群体中所获得的收益. 因此, 接下来只需计算这些表达式.

6.13 有限群体中的收益值

收益值由两部分组成: 即公共利益项 G 和惩罚项 S. 我们只考虑具有同级惩罚的 SR 型公共利益博弈.

在 PP 情形, 惩罚只影响惩罚者和背叛者. 如果有 Z 个惩罚者和 Y 个背叛者 $(Z + Y = M)$, 则前者须支付

$$S_{ZY} = \frac{\gamma Y(N-1)}{M-1}, \tag{6.90}$$

而后者须支付

$$S_{YZ} = \frac{\beta Z(N-1)}{M-1}. \tag{6.91}$$

此时, 所有其他的 S 项为 0.

如果群体只由 X 个合作者和 Y 个背叛者 $(X + Y = M)$ 组成, 则

$$G_{XY} = \sum_{k=0}^{N-1} H(k, N-1, X-1, M-1) \left(\frac{k+1}{N} r - 1 \right) c, \tag{6.92}$$

其中,

$$H(k, N-1, X-1, M-1) = \frac{\binom{X-1}{k}\binom{M-X}{N-k-1}}{\binom{M-1}{N-1}} \tag{6.93}$$

表示一个合作者在有 k 个其他合作者和 $N - 1 - k$ 个背叛者样本中的概率, 参看表达式 (6.81)①. 这样

$$G_{XY} = \frac{rc}{N} \left[1 + (X-1)\frac{N-1}{M-1} \right] - c. \tag{6.94}$$

①译注: 原文此处为 “(6.82)”.

类似的有

$$G_{YX} = \left(\frac{rc}{N}\right)\frac{N-1}{M-1}X, \tag{6.95}$$

$$G_{XZ} = G_{ZX} = (r-1)c, \tag{6.96}$$

$$G_{WX} = G_{WY} = G_{WZ} = \sigma, \tag{6.97}$$

$$G_{XW} = G_{ZW} = (r-1)c - \frac{W_{N-1}}{(M-1)_{N-1}}((r-1)c-\sigma), \tag{6.98}$$

$$G_{YW} = \frac{W_{N-1}}{(M-1)_{N-1}}\sigma, \tag{6.99}$$

$$G_{YZ} = \frac{Z(N-1)}{M-1}\left(\frac{rc}{N}\right), \tag{6.100}$$

以及

$$G_{ZY} = \frac{rc}{N}\left[\frac{(Z-1)(N-1)}{M-1}+1\right]-c. \tag{6.101}$$

现在可以计算两种情形的马尔可夫链的平稳分布, 即自愿参加和强制参加的情形. 前者, 惩罚者占优, 如图 6.5 所示, 而后者, 背叛者胜出, 如图 6.6 所示. 基于个体的模拟说明了这点. 如果参加公共利益博弈是自愿的而不是强制的, 那么基于惩罚的合作更容易实现, 如图 6.7 所示. 类似的结果也适用于其他情形 (即 OO 类型的公共利益博弈, 或是由 "惩罚机构" 提供的惩罚).

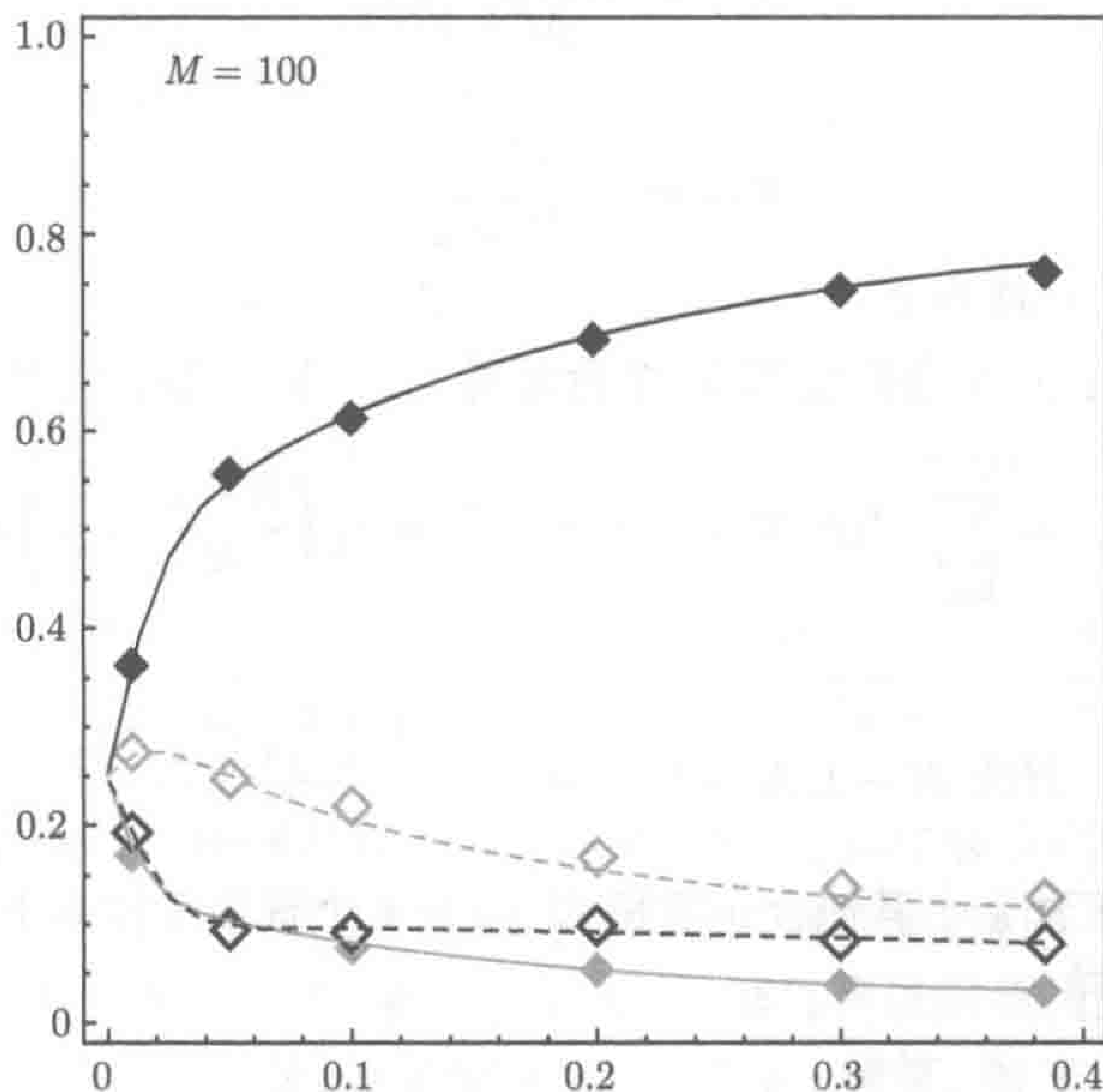


图 6.5 不同选择强度 s 对应的惩罚者, 合作者, 背叛者和自给自足参与者 (从上到下依此顺序) 的平稳分布. 参数 $M = 100$, $N = 5$, $r = 3$, $\sigma = 1$, $\beta = 1$, $\gamma = 0.3$, $c = 1$

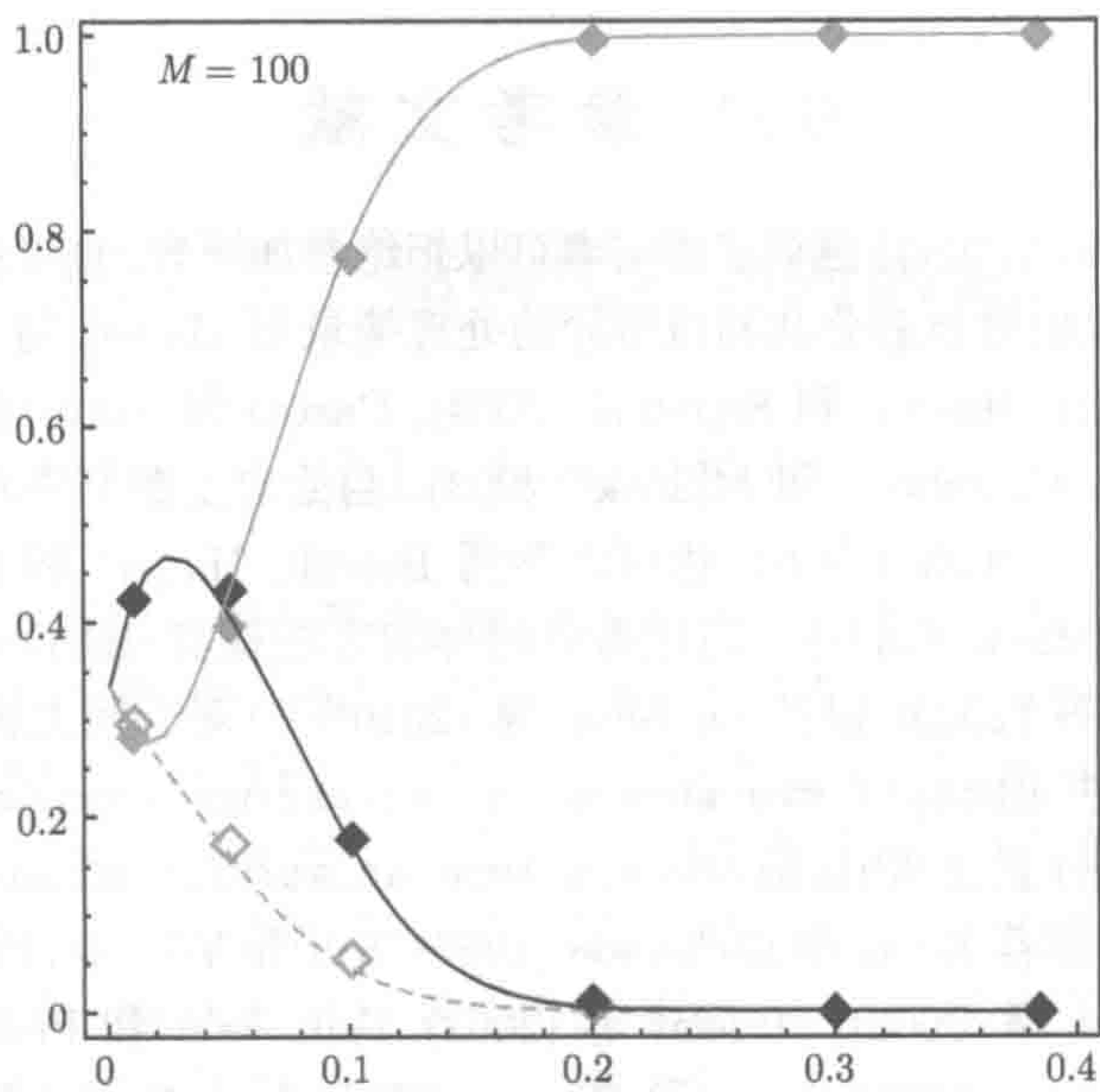


图 6.6 相同参数下强制博弈 (所有参与者必须参加) 的平稳分布. 现在的顺序是 (右边从上到下): 背叛者、惩罚者和合作者

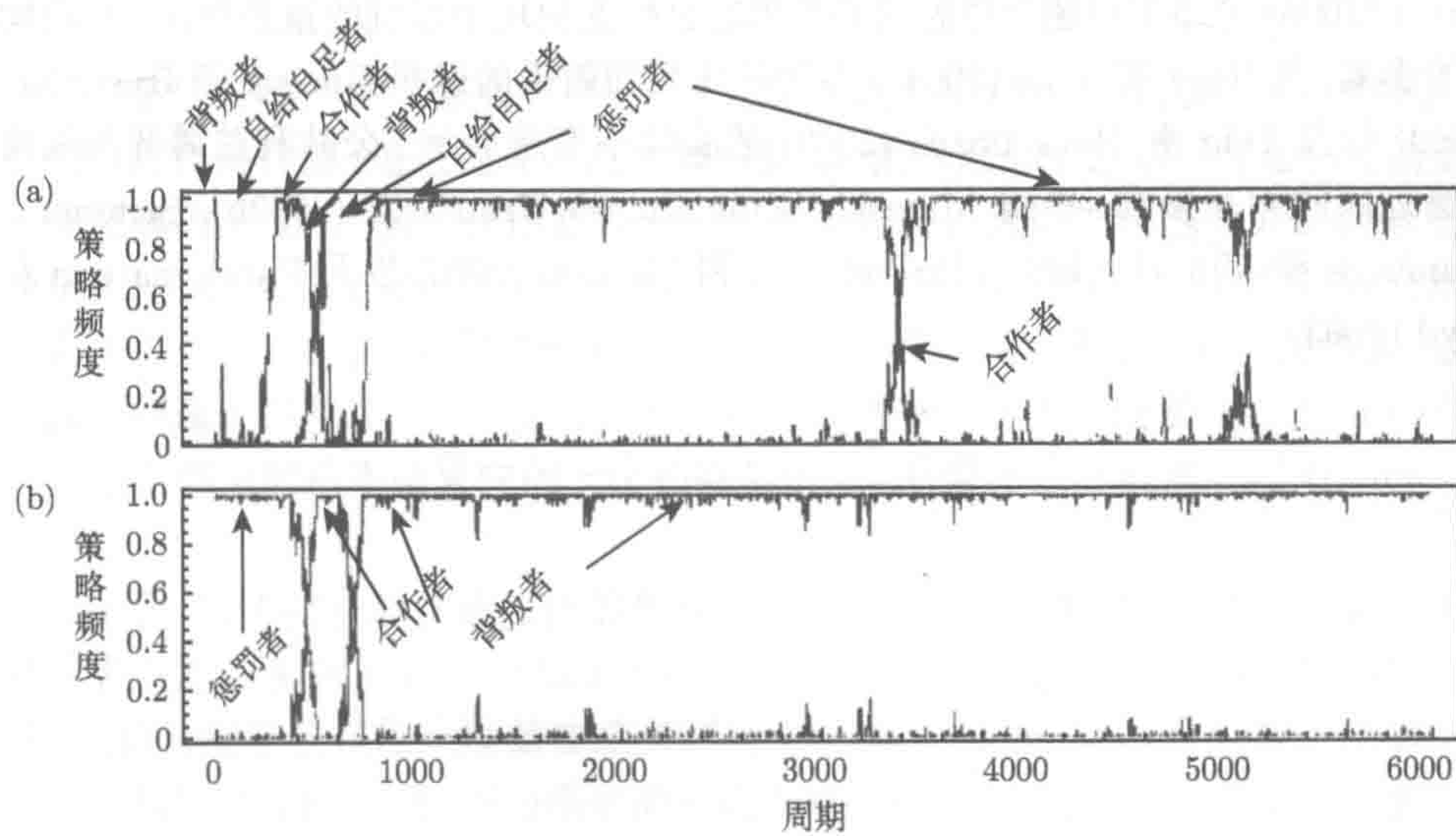


图 6.7 一种典型进化的简图 (a) 有惩罚的自愿参加的公共利益博弈, 及 (b) 有惩罚的强制参加的公共利益博弈. 对应的基于个体的计算机模拟 (Hauert, 2007)

6.14 参考文献

Orbell 和 Dawes (1993) 强调了参与者可以拒绝参加博弈, 也可参阅 Batali[①] 和 Kitcher (1995). 我们对自愿公共利益博弈的处理摘录自 Hauert 等 (2002a, 2002b), 也可以参阅 Hauert, Haiden 和 Sigmund (2004), Castro 和 Toro (2008), 实验性检验参阅 Semmann, Krambech 和 Milinski (2003). 自给自足参与者对惩罚出现的重要性最先由 Fowler (2005a) 提出, 也可以参阅 Brandt, Hauert 和 Sigmund (2006) 以及 Boyd 与 Mathew (2007). 通过有限群体的方法摘自 Hauert 等 (2007), 也可参阅 Traulsen 等 (2009) 以及 De Silva 等 (2010[②]). 对于网上模拟, 参阅 Hannelore De Silva 的主页`http://www.wu-wien.ac.at/usr/ma/hbrandt/publicgoods/` 和 Christoph Hauert 的主页`http://www.univie.ac.at/virtuallabs/`. 关于惩罚有大量的文献, 我们推荐 Boyd 和 Richerson (1992), Sell 和 Wilson (1999), Henrich 和 Boyd (2001), Boyd 等 (2003), Masclet 等 (2003), Falk, Fehr 和 Fischbacher (2005), Small 和 Loewenstein (2005), Xiao 和 Houser (2005) 以及 Dawes 等 (2007). 一些作者考虑二阶惩罚 (对不惩罚背叛者的给予惩罚), 但这对惩罚的出现影响很少, 似乎也几乎没有经验的支持, 见 Kiyonari 等 (2004) 以及 Fowler (2005b). Johnson 和 Bering (2006) 论述了对超自然惩罚的恐惧在合作的出现中起到的重要作用. 社会构架的影响, 见 Rege 和 Telle (2004). 各种产生负面效果的惩罚 Gneezy 和 Rustichini (2000) 以及 Fehr 和 Rockenbach (2003) 的实验中有所展示. 公共利益博弈与间接互惠之间的相互影响, 参阅 Milinski, Semmann 和 Krambeck (2002b), Semmann, Krambeck 和 Milinski (2004), Rochenbach 和 Milinski (2006) 以及 Panchanathan 和 Boyd (2003).

---

①译注: 原文此处为 "Battali".

②译注: 原文此处为 "2009".

第7章　结构化群体内的合作

7.1　结构化群体

到目前为止，我们总假设群体是完全融合的：正如群体遗传学常常假定随机交配，我们假定随机相遇．然而，这种假定通常是不现实的．一般来说，人类社会是高度结构化的，人们总是倾向于在家庭内部、邻里之间或者其他形式的社交圈以及团体内部进行交互．显然，这些结构化因素在合作的进化过程中发挥着重要的作用．

至此，我们的分析只是基于策略的视角，而完全忽略了社交网络的因素，尽管其重要性显而易见．最后这一章，我们将简要地讨论一些相关的内容，不是对此前的忽略作一个迟来的补救，而旨在指出一些主要的研究方向．

7.2　亲 缘 选 择

大量的合作产生于家庭内部．这是达尔文关于生存竞争理论的一个直接推论．那些促进自身传播（通过提高存活和繁殖能力）的基因必然比那些与之相反的基因拥有更高的频度．正如总是愿意帮助子女的父母在传递其遗传基因方面拥有明显的优势，兄弟姐妹之间的相互帮助也是如此．更准确地讲，帮助兄弟的基因有助于其自身的传播，因为该基因很可能为其兄弟所携带．这种“利己基因”的观点，通过亲缘选择理论的阐述，在过去的五十年中已经得到了极大的发展．此处，我们仅仅给出这套理论的基本结果中的一个初等实例，即所谓的*汉密尔顿法则* (Hamilton's rule).

下面，我们将从遗传适应性的角度来解释“收益”，即繁殖成功率．记 ρ 为两位参与者之间的*关联系数* (coefficient of relatedness)．这可以通过其它方式定义．此处，我们假定它仅仅用来度量血统关联性 (relatedness by descent)：即一个参与者携带的最新突变的基因（更为准确的术语为等位基因）正好被另一个参与者携带的概率．当然，如果追溯到人类始祖夏娃，任何两个人类个体总是相关的．但我们并不拥有完全相同的基因．发生在祖父身体上的突变将产生一个等位基因，该基因将有 1/2 的概率出现在子女身上，进而有 1/4 的概率出现在孙辈身上．在通常情况下（如果没有近亲繁殖），亲兄弟姐妹之间的关联系数为 1/2，而自己与侄子之间的关联系数则为 1/4，诸如此类，如图 7.1 所示．

从基因学的观点看, 这意味着亲属的繁殖成功也将促进我们自身的繁殖, 只是这种间接促进作用具有相关系数 $\rho < 1$. 考虑一个捐赠博弈, 其收益矩阵为

$$\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \tag{7.1}$$

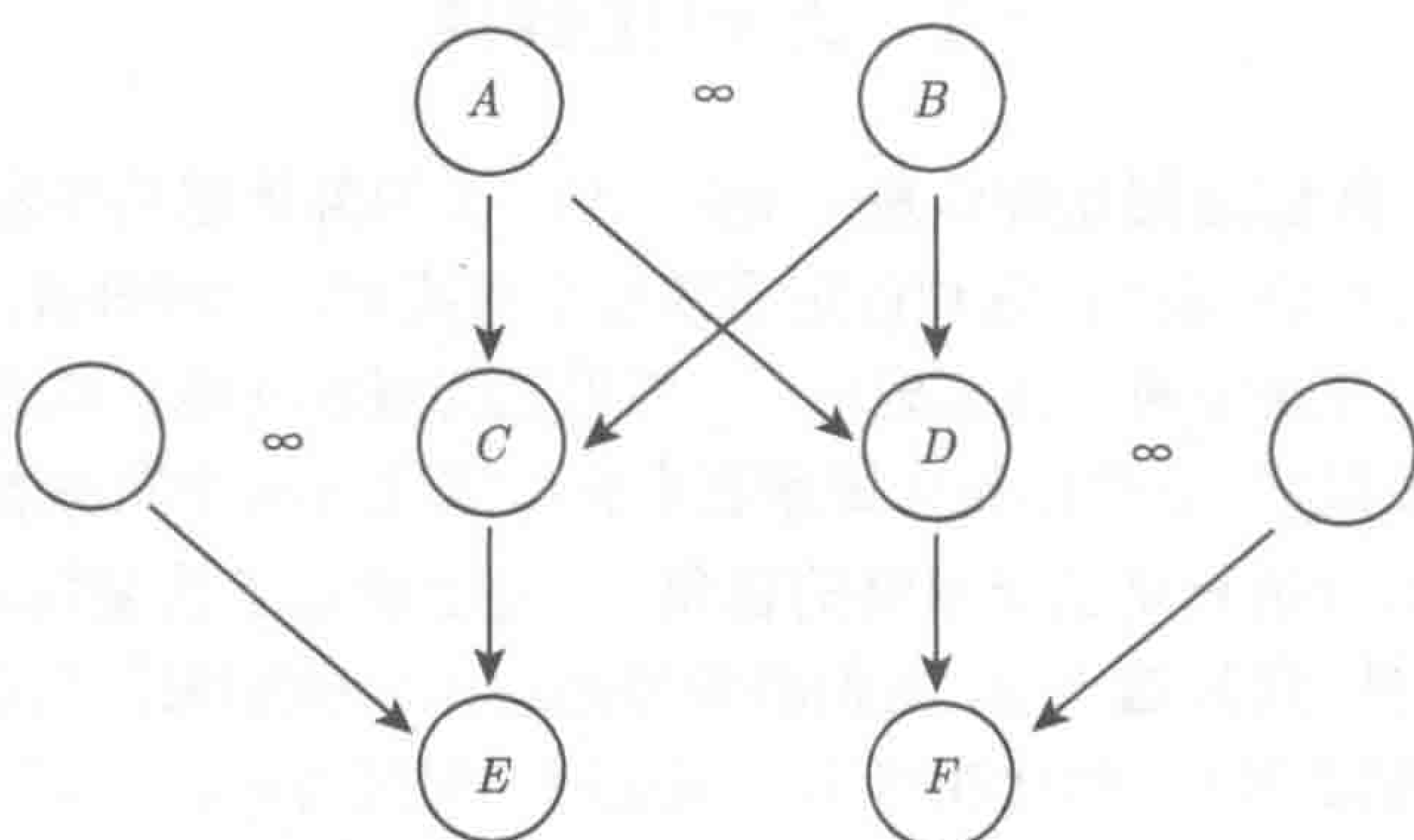


图 7.1 小型家庭的关联程度. A 和 B 为父母, C 和 D 为子女, E 和 F 为孙子. 单一个体中, 每个基因都有两份复制 (等位基因), 一份来自母亲, 一份来自父亲. A 的某个等位基因遗传给 C 的概率为 1/2. C 的某个等位基因来自 A 的概率也为 1/2. A 与 C 的关联度为 1/2. C 和 D (互为兄弟姐妹) 可能都继承了父辈的同一个突变等位基因, 或者来自 A (概率为 $1/2\times1/2=1/4$), 或者来自 B, 因此 C 和 D 的关联程度为 $1/4+1/4=1/2$. F 和 C (叔侄) 的关联程度为 $1/2\times1/2\times1/2+1/2\times1/2\times1/2=1/4$. E 和 F 的关联程度为 1/8

显然, 第二个策略优于第一个策略 (3.1). 但如果将对手的收益以因子 ρ 的比例附加到自己的收益上, 则收益矩阵就变为

$$\begin{pmatrix} (1+\rho)(b-c) & -c+b\rho \\ b-c\rho & 0 \end{pmatrix} \tag{7.2}$$

显然, 由条件 $\rho b > c$ 可推导出第一个策略 (即合作) 占优: 无论对手合作还是背叛, 选择合作总是最优的. 例如, 对于亲兄弟姐妹而言, 此条件意味着收益 (b) 必须超过成本 (c) 的两倍, 关系

$$\rho > \frac{c}{b} \tag{7.3}$$

称为汉密尔顿法则. 如果满足该法则, 则值得合作. 汉密尔顿是在群体遗传学背景下推导出该法则的. 很明显, 只要假设对手的收益将 (以某个因子 ρ 的比率) 作用到自己的收益上, 则该法则恒成立, 而无需将博弈限制在亲属之间. 该法则之所以

成立, 其原因很简单, 因为你喜欢你的对手. (当然一个标准的达尔文主义者可能要求给出这种情感的遗传学解释.)

前面我们曾得到与汉密尔顿法则相类似的不等式. 特别是, 直接回报理论中(3.4 节), 有重要的不等式

$$w > c/b, \tag{7.4}$$

其中, w 是新一轮发生的概率, 而对于间接回报理论 (4.3 节), 有不等式

$$q > c/b, \tag{7.5}$$

其中, q 为获悉对手在前一回合策略行为的概率.

7.3 网格博弈

现在考虑另一种情况. 假设参与者并非不停地随机移动, 而是长时间不动. 他们只与邻居进行交互. 考虑最简单的情形, 假设他们生活在一个巨大的棋盘上, 每一位参与者都位于网格点上. 于是每位参与者与其他八位参与者相邻, 并只与他们进行交互. 假设参与者必须参与收益矩阵为 (7.1) 的捐赠博弈, 且策略选择一致, 即在与八位邻居博弈时都选择 C(合作) 或者都选择 D(背叛). 背叛者的收益为 N_cb, 其中 $0 \leqslant N_c \leqslant 8$ 表示选择合作的邻居数量. 而合作者的收益为 $N_cb - 8c$. 假设参与者与八个邻居分别进行一轮捐赠博弈, 然后更新策略. 当某个邻居的收益最高时, 则选择该邻居的策略, 或者当自己的收益已经最高时, 则坚持自己的策略. 而当最高收益并列出现时, 则随机选择策略.

易见, 置身于一群背叛者包围圈中的孤立合作者将会消失. 而一个背叛者会扩散至其所有的八个邻居. 但是对于更大的集群, 最终结果可能依赖于具体的几何形状, 如图 7.2 所示. 例如, 有一条直线将背叛者和合作者分隔开, 那么处于前沿的合作者收益为 $5b - 8c$, 而处于前沿的背叛者收益为 $3b$. 因此, 如果 $b/c > 4$, 则合作者的收益总比背叛者好. 如果此条件满足, 那么前沿线将向前移动. 临近前沿线的第二列合作者的收益为 $8b - 8c$, 而如果 $b/c < 8/5$, 则背叛者将前移. 另一方面, 处于合作者方块内角落位置的参与者, 其收益为 $3b - 8c$, 而某些背叛者邻居的收益是 $3b$, 这个结果始终要好于处于角落的合作者. 但是, 处于背叛者方块内的角落位置的参与者, 其收益为 $5b$, 当 $b > 4c$, 这个值总是要比处于方块外与该背叛者对角相邻的合作者的收益值 $7b - 8c$ 小.

当然, 背叛者与合作者的位置不太可能按照一个简单的几何构型分布. 一般来讲, 不宜假设参与者位于一个规则的网格里, 比如一个无限大的棋盘. 相邻参与者的关系更为随机, 而由一个不规则的图形确定.

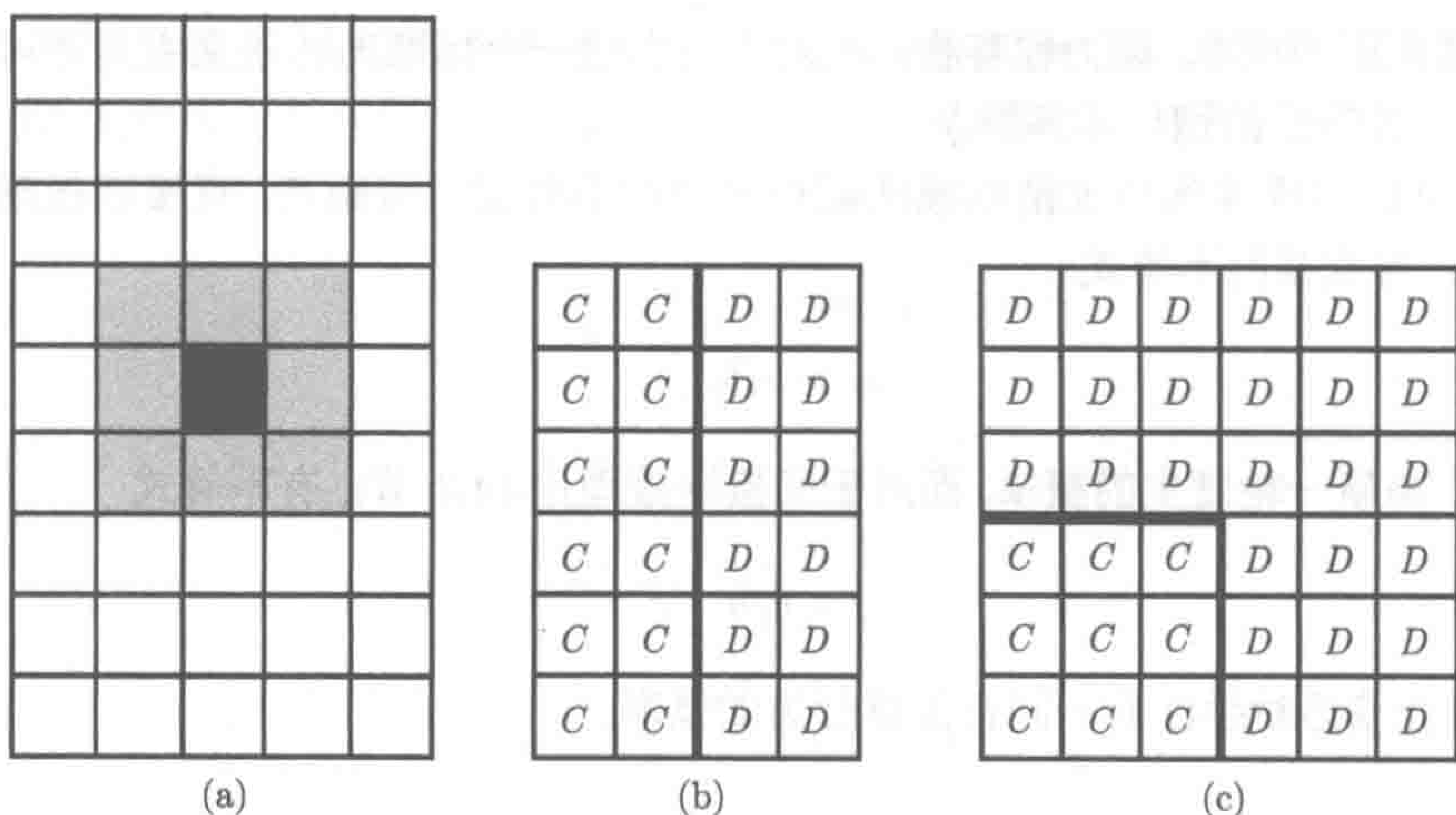


图 7.2　网格博弈: (a) 描述位于特定位置的参与者的邻居; (b) 背叛者与合作者之间的前沿分隔直线; (c) 描述被背叛者包围的位于角落的合作者方块

此外, 还有很多其他的建模方法可以用来描述策略是如何通过模仿或其他方式进行扩散的. 参与者可以异步更新策略, 也可以采用一些随机学习规则等. 但只要交互网络相对规则, 参与者模仿邻居的概率正比于邻居的适合度, 且该适合度只是弱依赖于收益自身, 根据经验可以判断, 合作总会获胜, 只要

$$k < \frac{b}{c}, \tag{7.6}$$

其中 k 是邻居的平均数量. 直观上, 这是合理的. 邻居越多, 独特性就越少, 也就意味着群体融合得越好.

7.4　Price 方程

只要群体被细分到亚群体, Price 方程就能提供一个用于计算进化的有效工具. 如图 7.3 所示, 假设亚群体 i ($1 \leqslant i \leqslant m$) 的个体数量为 N_i, 并且具有某个特征指标 p_i (例如, 在亚群体内合作者的频度, 或者全体成员的总质量), 则群体总规模为 $N = \sum N_i$, 而特征指标平均值为 $E(p) = \sum N_i p_i / N$.

假设个体繁殖, 因而下一代亚群体的规模变为 $N_i{}'$, 其特征指标为 $p_i{}'$. 于是特征指标的平均值变为 $E'(p') = \sum N_i{}' p_i{}' / N'$ (记号 E' 用以强调平均值决定于在新一代群体中亚群体所占的比重.) 分数 $N_i{}'/N_i$ 可以看作亚群体 i 的单位繁殖率, 或适合度, 记为 f_i. Price 方程建立了整个群体特征指标的平均增长量 $\Delta p = E'(p') - E(p)$ 与各亚群体特征指标的增长量 $\delta p_i = p_i' - p_i$ 之间的关系. 通常, 平均值的差异不同

于差异的平均值. 但是这两个表达式可以通过两个函数 $p: i \longmapsto p_i$ 与 $\boldsymbol{f}: i \longmapsto \boldsymbol{f}_i$ 的协方差进行关联, 即

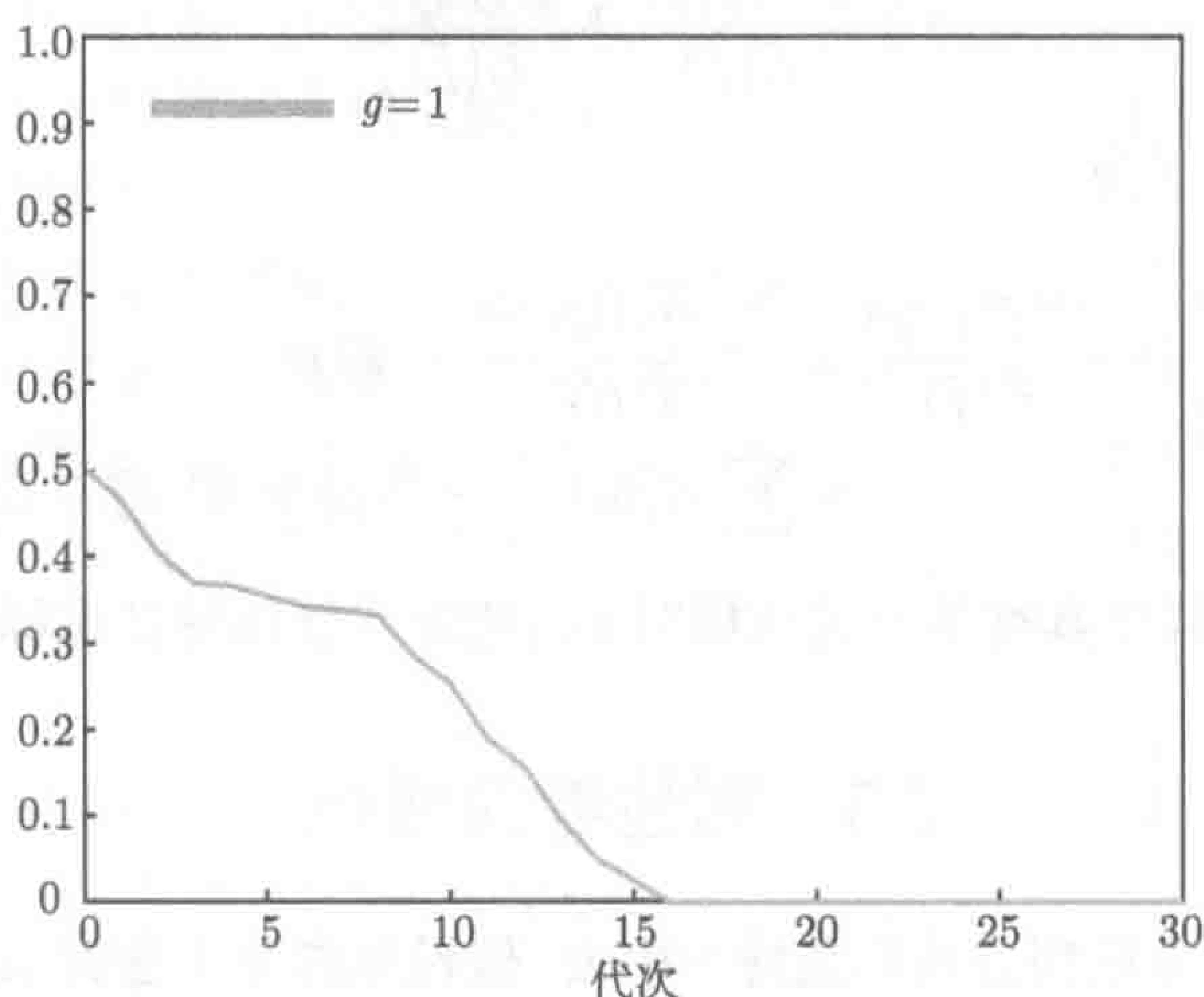


图 7.3 每一代亚群体都随机重组情形时合作者的频度

$$\mathrm{cov}(f,p) = E(fp) - E(f)E(p). \tag{7.7}$$

此处与前面相同, E 表示均值, 即

$$E(f) = \frac{1}{N}\sum N_i f_i, \tag{7.8}$$

$$E(fp) = \frac{1}{N}\sum N_i f_i p_i. \tag{7.9}$$

显然

$$E(f) = \frac{N'}{N}, \tag{7.10}$$

且

$$E(f\delta p) = \frac{1}{N}\sum N_i f_i \delta p_i = E(fp') - E(fp). \tag{7.11}$$

由于

$$E(fp') = \frac{1}{N}\sum f_i p_i' N_i = \frac{1}{N}\sum \frac{N_i'}{N_i} p_i' N_i = \frac{N'}{N}\sum \frac{N_i' p_i'}{N'} = E(f)E'(p'), \tag{7.12}$$

由协方差表达式 (7.7) 可得 Price 方程

$$E(f)[E'(p') - E(p)] = \mathrm{cov}(f,p) + E(f\delta p) \tag{7.13}$$

或者, 若 $N' > 0$, 即若 $E(f) \neq 0$,

$$\Delta p = \frac{\text{cov}(f,p)}{E(f)} + \frac{E(f\delta p)}{E(f)}. \tag{7.14}$$

其中 $\text{cov}(f,p)/E(f)$ 为

$$\begin{aligned}\frac{E(fp)}{E(f)} - \frac{E(f)E(p)}{E(f)} &= \frac{\sum N_i f_i p_i/N}{N'/N} - E(p)\\ &= \sum N_i' p_i/N' - E(p) = E'(p) - E(p).\end{aligned} \tag{7.15}$$

$E'(p)$ 可以看作当群体具有下一代规模时 p_i (初始代群体特征指标) 的平均值.

7.5　利他者的增长

将 Price 方程应用到公共利益博弈中去, 假设亚群体 i 含有 X_i 个合作者及 Y_i 个背叛者, 且 $X_i + Y_i = N_i$. 合作者的频度为 $p_i = X_i/N_i$. 假设在每个亚群体内部进行公共利益博弈. 如果是 SR 类型 (6.2 节), 则背叛者收益 P_d 及合作者收益 P_c 分别为

$$P_d(i) = rcX_i/N_i = rcp_i \quad (r > 1) \tag{7.16}$$

和

$$P_c(i) = rcp_i - c. \tag{7.17}$$

令适合度函数为基准项 $B > 0$ (对所有的亚群体 i) 与博弈中收益项的凸组合. 则亚群体 i 中背叛者和合作者的适合度分别为 $(1-s)B + sP_d(i)$ 及 $(1-s)B + sP_c(i)$, 其中 $s \in [0,1]$ 为选择系数 (s 的主要作用是确保适合度为正, 这可以通过取足够小的 s 来保证). 亚群体 i 的平均适合度为

$$f_i = (1-s)B + s[rcp_i(p_i + (1-p_i)) - cp_i] = (1-s)B + sc(r-1)p_i. \tag{7.18}$$

Price 方程 (7.13) 的右端是两项之和, 第二项非正. 实际上, 在每个亚群体内, 背叛者的收益都好于合作者, 因此, 合作者的期望频度将下降: $\delta p_i \leqslant 0$ (对所有的 i). 但第一项是非负的. 实际上

$$\text{cov}(f,p) = sc(r-1)\text{cov}(p,p) = sc(r-1)\text{var}(p) \geqslant 0. \tag{7.19}$$

因此, 如果 p 的方差足够大, 即合作者在各亚群体中的比例存在足够大的差异, 则在整个群体中合作者的平均适合度逐代增长. 此外, 该结果也适用于公共利益博弈

中 OO 的情形. 当背叛者在每个亚群体内扩散时, 那些拥有较少背叛者的亚群体发展得更为迅速, 并由此获得补偿.

需要强调的是, 这是一个短暂的效应. 随着代际更替, 亚群体将变得越来越同质化 (因为每一个亚群体终将越来越由背叛者主导). 因此, 方差将逐步变小, 而合作者不再增加, 如图 7.3 所示.

然而, 可以有很多方式来避免这一现象: 例如, 通过允许亚群体之间适当的迁移. 另一种方式是每隔几代将亚群体随机重组: 如果在一定范围内巧妙的选择间隔世代数, 方差会不断地反弹, 如图 7.4 所示. 也可以假设合作者倾向于团结其他合作者. 当然, 背叛者也更愿意与合作者联合, 但是只要合作者不接纳, 背叛者将被拒之门外.

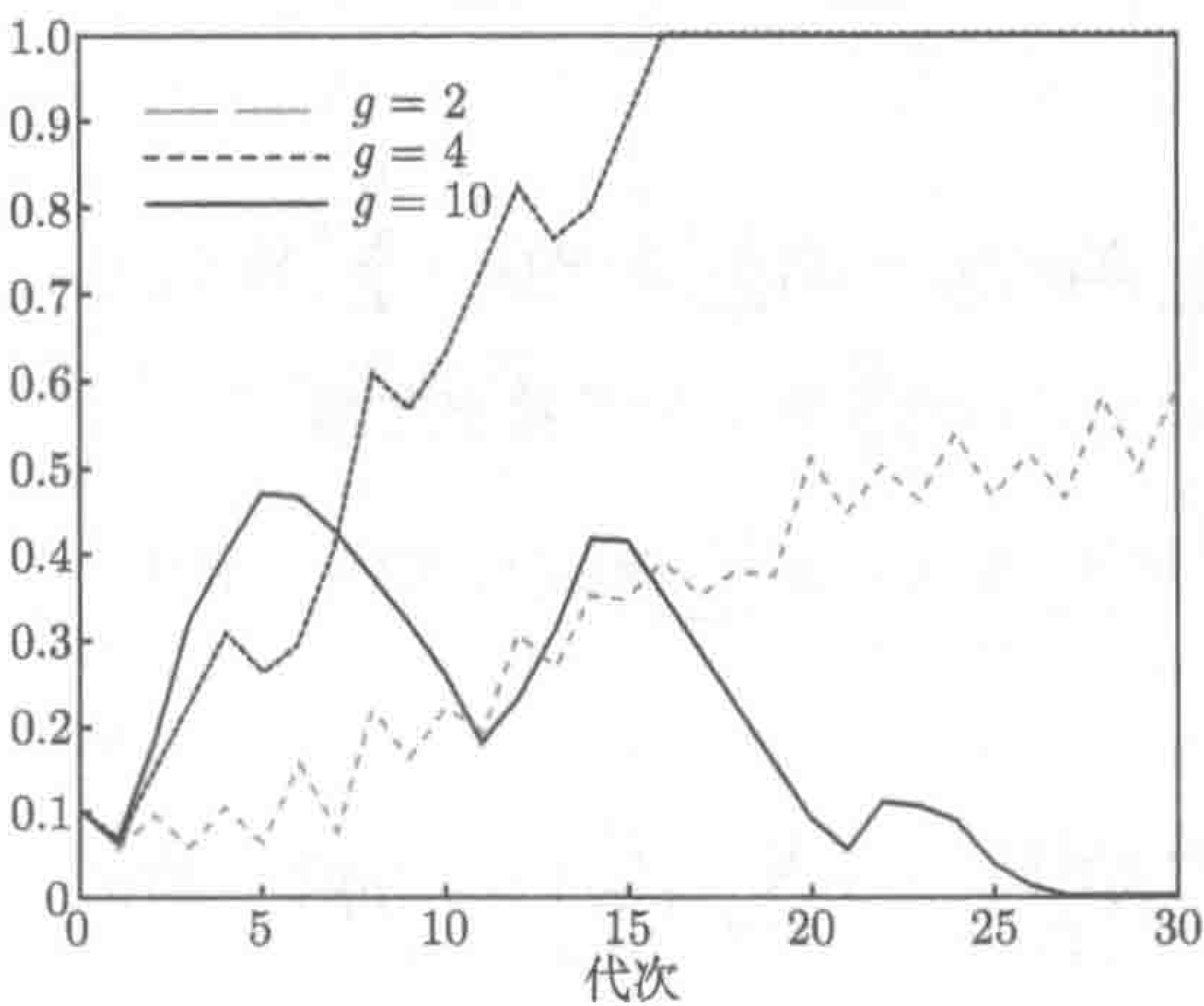


图 7.4 基于 Fletcher 及 Zwick (2004) 和 Pflügl (2007) 的研究, 每隔二代、四代及十代之后随机分组情况下合作者的频度. 重组的世代数间隔 g 在合作进化过程中起着至关重要的作用

另一个现实但更为极端的方式是去掉那些拥有过多背叛者的亚群体, 代之以从合作者居多的亚群体中随机分割出来的新的亚群体. 例如, 可以假设背叛者居多的亚群体在战争中被那些勇于牺牲的利他主义者亚群体所消灭. 这种方法可以追溯到达尔文, 他曾写到: "毫无疑问, 由乐于帮助并愿为部落利益牺牲的成员所构成的部落会胜出, 这就是自然选择".

7.6 另一种汉密尔顿法则

关于 p 的方差的相关解释可以通过寻求合作者频度增加的充分必要条件来给出. 分别记 x_i' 和 y_i' 为合作者频度 X_i' 及背叛者频度 Y_i' 在下一代的期望值. 由于

$$x_i' = X_i[(1-s)B + sP_c(i)], \tag{7.20}$$

和

$$y_i' = Y_i[(1-s)B + sP_d(i)], \tag{7.21}$$

(利用 $x' = \sum x_i'$, $y' = \sum y_i'$ 以及 $n' = x' + y'$), 条件

$$x'/n' > X/N \tag{7.22}$$

转化为

$$\frac{(1-s)BX + s\sum X_i P_c(i)}{(1-s)BN + s(\sum X_i P_c(i) + \sum Y_i P_d(i))} > \frac{X}{N}. \tag{7.23}$$

这又蕴含了

$$N\sum X_i P_c(i) > X(\sum X_i P_c(i) + \sum Y_i P_d(i)). \tag{7.24}$$

利用方程 (7.16) 和 (7.17), 并将含有 r 的项移至左端

$$r\left[N\sum X_i p_i - X\sum N_i p_i] > X(N-X)\right], \tag{7.25}$$

两端除以 N^2 可得

$$r\left[\sum p_i \frac{X_i}{N} - E(p)\sum p_i \frac{N_i}{N}\right] > E(p)(1-E(p)). \tag{7.26}$$

r 的因式可以写成

$$\frac{1}{N}\sum {p_i}^2 N_i - E(p)\frac{1}{N}\sum N_i p_i = E(p^2) - (E(p))^2 = var(p). \tag{7.27}$$

因此, 合作者的占比增加当且仅当

$$r[\mathrm{var}(p)] > E(p)(1-E(p)). \tag{7.28}$$

将右端记为 $\mathrm{Var}(p)$. 它是一个随机变量的方差, 该变量通过在整个群体中随机抽取单一个体并检测其是合作 (值 1) 或背叛 (值 0) 而得到. 注意到, c 为个体对公共利益做出贡献的成本, 而 $rc =: b$ 为整个亚群体的累计收益, 则有

$$\frac{\mathrm{var}(p)}{\mathrm{Var}(p)} > \frac{c}{b}, \tag{7.29}$$

这就是另一个汉密尔顿型不等式. 它凸显了 p 的方差的重要性. 如果合作者的比例在各个亚群体中都是相似的, 则合作者在整个群体中的比例将下降.

7.7 参考文献

Nowak (2006b) 给出了研究结构化和非结构化群体内的合作的统一方法. 亲缘选择理论的鼻祖是 W. D. Hamilton, 参见其开创性的论文的重印版 Hamilton (1996) 以及 Grafen (1984) 的清晰阐述. 关于该理论的现代描述, 参见 Frank (1998), Taylor 和 Frank (1996), Pepper (2000), Taylor, Day 和 Wild (2007), Skyrms 和 Pemantle (2000), West, Pen 和 Griffin (2002), Rousset (2004) 以及 Lehmann 和 Keller (2006). 网格博弈由 Nowak 和 May (1992) 提出, 参见 Nowak, Bonhoeffer 和 May (1994), Lindgren 和 Nordahl (1994) 以及 Hauert 和 Szabó (2003). Schelling (1971) 以及 Durrett 和 Levin (1994) 指出了空间关系的重要性. Hutson 和 Vickers (1995) 以及 Ferrière 和 Michod (1996) 研究了囚徒困境博弈的反应扩散方程. Lieberman①, Hauert 和 Nowak(2005), Szabó 和 Fáth (2007), Santos, Pacheco 和 Lenaerts (2006) 以及 Pacheco 等 (2008) 研究了不规则网络和图形位置的合作博弈. Nowak 和 Sigmund (2000) 研究了空间的最后通牒博弈; Brandt 等 (2007) 研究了空间的间接回报; Hauert 和 Doebeli (2004) 以及 Doebeli, Hauert 和 Killingback (2004) 研究了空间的雪堆博弈; Brandt, Hauert 和 Sigmund (2003) 以及 Nakamaru 和 Iwasa (2005, 2006) 研究了空间的公共利益博弈; Szabó和 Hauert (2002) 以及 Szabó和 Vukov (2004) 研究了自愿参加所起的作用. Price (1970) 首次提出 Price 方程, 也可参见 Grafen (2000), Page 和 Nowak (2002) 描述了其与复制方程之间有趣的联系. Price 方程已经成为许多研究工作的基础, 例如, Frank (1998), Sober 和 Wilson (1998) 或者 Gintis (2000). 7.5 节和 7.6 节的内容以及图 7.3 是来自 Fletcher 和 Zwick (2004, 2007) 的工作. Cohen 和 Eshel (1976), Bergstrom (2002) 以及 Killingback, Bieri 和 Flatt (2006) 采用不同的方法研究了群体选择问题. Lehmann 和 Keller (2006) 以及 Lehmann 等 (2007) 说明了群体选择是亲缘选择, 也可参见 West, Griffin 和 Gardner (2007).

①译注: 原文此处为 “Liebermann”.

参考文献

Abreu, D., and A. Rubinstein. 1988. "The structure of Nash equilibrium in repeated games with finite automata." *Econometrica* 56:1259–1281.

Alexander, R. D. 1987. *The Biology of Moral Systems*. New York: Aldine de Gruyter.

Andreoni, J., W. Harbaugh, and L. Vesterlund. 2003. "The carrot or the stick: Rewards, punishments, and cooperation." *American Economic Review* 93:893–902.

Ashraf, N., I. Bohnet, and N. Piankov. 2006. "Decomposing trust and trustworthiness." *Experimental Economics* 9:193–208.

Aumann, R. J. 1981. "Survey of repeated games." In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern, Gesellschaft, Recht, Wirtschaft*, vol. 4, ed. V. Bohm, 11–42. Mannheim: Bibliographisches Institut.

Axelrod, R. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, R., and D. Dion. 1988. "The further evolution of cooperation." *Science* 242:1385–1390.

Axelrod, R., and W. D. Hamilton. 1981. "The evolution of cooperation." *Science* 211:1390–1396.

Banks, J. S., and R. K. Sundaram. 1990. "Repeated games, finite automata, and complexity." *Games and Economic Behavior* 2:97–117.

Batali, J., and P. Kitcher. 1995. "Evolution of altriusm in optional and compulsory games." *Journal of Theoretical Biology* 175:161–71.

Bateson, M., D. Nettle, and G. Roberts. 2006. "Cues of being watched enhance cooperation in a real-world setting." *Biology Letters* 2:412–414.

Baumeister, R. F., E. Bratslavsky, C. Finkenauer, and K. D. Vohs. 2001. "Bad is stronger than good." *Review of General Psychology* 5:323–70.

Bendor, J., and P. Swistak. 1995. "Types of evolutionary stability and the problem of cooperation." *Proceedings of the National Academy of Sciences* 92:3596–600.

Bendor, J., and P. Swistak. 2001. "The evolution of norms." *American Journal of Sociology* 106:1493–545.

Berg, J., J. Dickhaut, and K. McCabe. 1995. "Trust, reciprocity, and social history." *Games and Economic Behavior* 10:122–42.

Bergstrom, T. C. 2002. "Evolution of social behavior: Individual and group selection." *Journal of Economic Perspectives* 16:67–88.

Binmore, K. G. 1994. *Playing Fair: Game Theory and the Social Contract*. Cambridge,

MA: MIT Press.

Binmore, K. G., and L. Samuelson. 1992. "Evolutionary stability in repeated games played by finite automata." *Journal of Economic Theory* 57:278–305.

Boerlijst, M. C., M. A. Nowak, and K. Sigmund. 1997a. "The logic of contrition." *Journal of Theoretical Biology* 185:281–93.

Boerlijst, M. C., M. A. Nowak, and K. Sigmund. 1997b. "Equal pay for all prisoners." *The American Mathematical Monthly* 104:303–5.

Bohnet, I., and R. Croson. 2004. "Trust and trustworthiness." *Journal of Economic Behavior & Organization* 55:443–45.

Bolton, G. E., E. Katok, and A. Ockenfels. 2004. "How effective are electronic reputation mechanisms? An experimental investigation." *Management Science* 50:1587–602.

Bolton, G. E., E. Katok, and A. Ockenfels. 2005. "Cooperation among strangers with limited information about reputation." *Journal of Public Economics* 89:1457–68.

Bolton, G. E., and A. Ockenfels. 2000. "Erc: A theory of equity, reciprocity, and competition." *The American Economic Review* 90:166–93.

Bomze, I. M. 1983. "Lotka-Volterra equation and replicator dynamics: A two-dimensional classification." *Biological Cybernetics* 48:201–11.

Bowles, S., and H. Gintis. 2002. "Homo reciprocans." *Nature* 415:125–28.

Boyd, R. 1989. "Mistakes allow evolutionary stability in the repeated Prisoner's Dilemma game." *Journal of Theoretical Biology* 136:47–56.

Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson. 2003. "The evolution of altruistic punishment." *Proceedings of the National Academy of Sciences* 100:3531–35.

Boyd, R., and S. Mathew. 2007. "A narrow road to cooperation." *Science* 316:1858–59.

Boyd, R., and P. J. Richerson. 1989. "The evolution of indirect reciprocity." *Social Networks* 11:213–36.

Boyd, R., and P. J. Richerson. 1992. "Punishment allows the evolution of cooperation (or anything else) in sizable groups." *Ethology and Sociobiology* 13:171–95.

Brandt, H., C. Hauert, and K. Sigmund. 2003. "Punishment and reputation in spatial public goods games." *Proceedings of the Royal Society B* 270:1099–104.

Brandt, H., C. Hauert, and K. Sigmund. 2006. "Punishing and abstaining for public goods." *Proceedings of the National Academy of Sciences* 103:495–97.

Brandt, H., H. Ohtsuki, Y. Iwasa, and K. Sigmund. 2007. "A survey of indirect reciprocity." In *Mathematics for Ecology and Environmental Sciences*, eds. Y. Takeuchi, Y. Iwasa, and K. Sato, 21–49. Berlin: Springer.

Brandt, H., and K. Sigmund. 2004. "The logic of reprobation: assessment and action rules for indirect reciprocation." *Journal of Theoretical Biology* 231:475–86.

Brandt, H., and K. Sigmund. 2005. "Indirect reciprocity, image scoring, and moral hazard." *Proceedings of the National Academy of Sciences* 102:2666–70.

Brandt, H., and K. Sigmund. 2006. "The good, the bad and the discriminator: Errors in direct and indirect reciprocity." *Journal of Theoretical Biology* 239:183–94.

Brosnan, S. F., and F. B. M. de Waal. 2003. "Monkeys reject unequal pay." *Nature* 425:297–99.

Brown, D. E. 1991. *Human Universals.* New York: McGraw-Hill.

Bshary, R., and A. S. Grutter. 2005. "Punishment and partner switching cause cooperative behaviour in a cleaning mutualism." *Biology Letters* 1:396–99.

Bshary, R., and A. S. Grutter. 2006. "Image scoring and cooperation in a cleaner fish mutualism." *Nature* 441:975–78.

Burnham, T., and B. Hare. 2007. "Engineering human cooperation: Does involuntary neural activation increase public goods contributions?" *Human Nature* 18:88–108.

Burnham, T. C., and D. D. P. Johnson. 2005. "The biological and evolutionary logic of human cooperation." *Analyse & Kritik* 27:113–35.

Camerer, C. F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton: Princeton University Press.

Camerer, C. F., and E. Fehr. 2006. "When does 'economic man' dominate social behavior?" *Science* 311:47–52.

Carpenter, J. P., G. W. Harrison, and J. A. List, eds. 2005. *Field Experiments in Economics, Research in Experimental Economics*, vol. 10. Amsterdam: Elsevier.

Castro, L., and M. A. Toro. 2008. "Iterated Prisoner's Dilemma in an asocial world dominated by loners, not by defectors." *Theoretical Population Biology* 74:1–5.

Chalub, F. A. C. C., F. C. Santos, and J. M. Pacheco. 2006. "The evolution of norms." *Journal of Theoretical Biology* 241:233–40.

Charness, G., and M. Dufwenberg. 2006. "Promises and partnership." *Econometrica* 74:1579–601.

Charness, G., and E. Haruvy. 2002. "Altruism, equity, and reciprocity in a gift-exchange experiment: An encompassing approach." *Games and Economic Behavior* 40:203–31.

Clutton-Brock, T. H., and G. A. Parker. 1995. "Punishment in animal societies." *Nature* 373:209–16.

Cohen, D., and I. Eshel. 1976. "On the founder effect and the evolution of altruistic traits." *Theoretical Population Biology* 10:276–302.

Colman, A. M. 1995. *Game Theory and its Applications in the Social and Biological Sciences.* International Series in Social Psychology, Oxford: Butterworth-Heinemann.

Colman, A. M. 2006. "The puzzle of cooperation." *Nature* 440:744–45.

Cose, E. 2004. *Bone to Pick: Of Forgiveness, Reconciliation, Reparation and Revenge.* New York: Atria Books.

Cox, J. C. 2004. "How to identify trust and reciprocity." *Games and Economic Behavior* 46:260–81.

Cressman, R. 2003. *Evolutionary Dynamics and Extensive Form Games, Economic Learning and Social Evolution*, vol. 5. Cambridge, MA: MIT Press.

Cross, J. G., and M. J. Guyer. 1980. *Social Traps.* Ann Arbor: University of Michigan Press.

Dawes, C. T., J. H. Fowler, T. Johnson, R. McElreath, and O. Smirnov. 2007. "Egalitarian motives in humans." *Nature* 446:794–96.

Dawes, R. M. 1980. "Social dilemmas." *Annual Review of Psychology* 31:169–93.

Dawkins, R. 1989. *The Selfish Gene.* 2nd ed. Oxford: Oxford University Press.

de Quervain, D. J.-F., U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, and E. Fehr. 2004. "The neural basis of altruistic punishment." *Science* 305:1254–58.

De Silva, H., C. Hauert, A. Traulsen, and K. Sigmund. 2010. "Freedom, enforcement, and the social dilemma of strong altruism." *Journal of Evolutionary Economics* 20:203–17.

de Waal, F. B. M. 1996. *Good Natured: The Origins of Right and Wrong in Humans and Other Animals.* Cambridge, MA: Harvard University Press.

DeBruine, L. M. 2002. "Facial resemblance enhances trust." *Proceedings of the Royal Society B* 269:1307–12.

Dickinson, D. L. 2001. "The carrot vs. the stick in work team motivation." *Experimental Economics* 4:107–24.

Dieckmann, U., and R. Law. 1996. "The dynamical theory of coevolution: a derivation from stochastic ecological processes." *Journal of Mathematical Biology* 34:579–612.

Dieckmann, U., and J. A. J. Metz, eds. 2009. *Elements of Adaptive Dynamics.* Cambridge: Cambridge University Press.

Doebeli, M., C. Hauert, and T. Killingback. 2004. "The evolutionary origin of cooperators and defectors." *Science* 306:859–62.

Duwfenberg, M., U. Gneezy, W. Güth, and E. van Damme. 2001. "Direct vs. indirect reciprocity: An experiment." *Homo Oeconomicus* 18:19–30.

Dugatkin, L. A. 1997. *Cooperation Among Animals: An Evolutionary Perspective.* Oxford Series in Ecology and Evolution, New York: Oxford University Press.

Dunbar, R. 1996. *Grooming, Gossip and the Evolution of Language.* Cambridge, MA: Harvard University Press.

Durrett, R., and S. A. Levin. 1994. "The importance of being discrete (and spatial)." *Theoretical Population Biology* 46:363–94.

Eckel, C. C., and R. K. Wilson. 2004. "Is trust a risky decision?" *Journal of Economic Behavior & Organization* 55:447–65.

Ellison, G. 1994. "Cooperation in the Prisoner's Dilemma with anonymous random matching." *The Review of Economic Studies* 61:567–88.

Engelmann, D., and U. Fischbacher. 2002. "Indirect reciprocity and strategic reputation building in an experimental helping game." Working paper, University of Zürich.

Engelmann, D., and U. Fischbacher. 2009. "Indirect reciprocity and strategic reputation building in an experimental helping game." *Games and Economic Behavior* 67:399–407.

Falk, A., E. Fehr, and U. Fischbacher. 2005. "Driving forces behind informal sanctions." *Econometrica* 73:2017–30.

Fehr, E., and U. Fischbacher. 2003. "The nature of human altruism." *Nature* 425:785–91.

Fehr, E., and U. Fischbacher. 2004. "Third-party punishment and social norms." *Evolution and Human Behavior* 25:63–87.

Fehr, E., and S. Gächter. 2000. "Cooperation and punishment in public goods experiments." *American Economic Review* 90:980–94.

Fehr, E., and S. Gächter. 2002. "Altruistic punishment in humans." *Nature* 415:137–40.

Fehr, E., and J. Henrich. 2003. "Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism." In *Genetic and Cultural Evolution of Cooperation*, ed. P. Hammerstein, 55–82. Cambridge, MA: MIT Press.

Fehr, E., and B. Rockenbach. 2003. "Detrimental effects of sanctions on human altruism." *Nature* 422:137–40.

Fehr, E., and K. M. Schmidt. 1999. "A theory of fairness, competition, and cooperation." *The Quarterly Journal of Economics* 114:817–68.

Ferrière, R. 1998. "Help and you shall be helped." *Nature* 393:517–19.

Ferrière, R., and R. E. Michod. 1996. "The evolution of cooperation in spatially heterogeneous populations." *The American Naturalist* 147:692–717.

Fishman, M. A. 2003. "Indirect reciprocity among imperfect individuals." *Journal of Theoretical Biology* 225:285–92.

Fishman, M. A., A. Lotem, and L. Stone. 2001. "Heterogeneity stabilizes reciprocal altruism interactions." *Journal of Theoretical Biology* 209:87–95.

Fletcher, J. A., and M. Zwick. 2004. "Strong altruism can evolve in randomly formed groups." *Journal of Theoretical Biology* 228:303–13.

Fletcher, J. A., and M. Zwick. 2007. "The evolution of altruism: Game theory in multilevel selection and inclusive fitness." *Journal of Theoretical Biology* 245:26–36.

Fowler, J. H. 2005a. "Human cooperation: Second-order free-riding problem solved?" *Nature* 437:E8.

Fowler, J. H. 2005b. "Altruistic punishment and the origin of cooperation." *Proceedings of the National Academy of Sciences* 102:7047–49.

Frank, R. H. 1988. *Passions within Reason: The Strategic Role of the Emotions.* New York: W. W. Norton & Company.

Frank, S. A. 1998. *Foundations of Social Evolution.* Princeton: Princeton University Press.

Frank, S. A. 2003. "Perspective: Repression of competition and the evolution of coopera-

tion." *Evolution* 57:693–705.

Frean, M. R. 1994. "The prisoner's dilemma without synchrony." *Proceedings of the Royal Society B* 257:75–79.

Fudenberg, D., and D. K. Levine. 1998. *The Theory of Learning in Games, Economic Learning and Social Evolution*, vol. 2. Cambridge, MA: MIT Press.

Fudenberg, D., and E. Maskin. 1986. "The folk theorem in repeated games with discounting or with incomplete information." *Econometrica* 54:533–54.

Fudenberg, D., and E. Maskin. 1990. "Evolution and cooperation in noisy repeated games." *The American Economic Review* 80:274–79.

Gardner, A., and S. A. West. 2004. "Cooperation and punishment, especially in humans." *The American Naturalist* 164:753–64.

Gaunersdorfer, A., J. Hofbauer, and K. Sigmund. 1991. "On the dynamics of asymmetric games." *Theoretical Population Biology* 39:345–57.

Gintis, H. 2000. *Game Theory Evolving.* Princeton: Princeton University Press.

Gintis, H., S. Bowles, R. Boyd, and E. Fehr. 2003. "Explaining altruistic behavior in humans." *Evolution and Human Behavior* 24:153–72.

Gintis, H., S. Bowles, R. Boyd, and E. Fehr, eds. 2005. *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life.* Cambridge, MA: MIT Press.

Gneezy, U., and A. Rustichini. 2000. "A fine is a price." *The Journal of Legal Studies* 29:1–17.

Grafen, A. 1984. "Natural selection, kin selection and group selection." In *Behavioral Ecology*, eds. J. R. Krebs and N. B. Davies, 2nd ed., 62–84. Oxford: Blackwell Scientific Publications.

Grafen, A. 2000. "Developments of the Price equation and natural selection under uncertainty." *Proceedings of the Royal Society B* 267:1223–27.

Gürerk, O., B. Irlenbusch, and B. Rockenbach. 2006. "The competitive advantage of sanctioning institutions." *Science* 312:108–11.

Güth, W., M. Königstein, N. Marchand, and K. Nehring. 2001. "Trust and reciprocity in the investment game with indirect reward." *Homo Oeconomicus* 18:241–62.

Güth, W., R. Schmittberger, and B. Schwarze. 1982. "An experimental analysis of ultimatum bargaining." *Journal of Economic Behavior & Organization* 3:367–88.

Haidt, J. 2007. "The new synthesis in moral psychology." *Science* 316:998–1002.

Haley, K. J., and D. M. Fessler. 2005. "Nobody's watching? Subtle cues affect generosity in an anonymous economic game." *Evolution and Human Behavior* 26:245–56.

Hamilton, W. D. 1996. *Evolution of Social Behaviour, Narrow Roads of Gene Land: The Collected Papers of W. D. Hamilton*, vol. 1. New York: W. H. Freeman.

Hammerstein, P., ed. 2003. *Genetic and Cultural Evolution of Cooperation.* Cambridge,

MA: MIT Press.

Harbaugh, W. T. 1998. "The prestige motive for making charitable transfers." *The American Economic Review* 88:277–82.

Hardin, G. 1968. "The tragedy of the commons." *Science* 162:1243–48.

Härdling, R. 2007. "Fairness evolution in the Ultimatum game is a function of reward size." *Journal of Theoretical Biology* 246:720–24.

Hauert, C., S. De Monte, J. Hofbauer, and K. Sigmund. 2002a. "Volunteering as Red Queen mechanism for cooperation in public goods games." *Science* 296:1129–32.

Hauert, C., S. De Monte, J. Hofbauer, and K. Sigmund. 2002b. "Replicator dynamics for optional public good games." *Journal of Theoretical Biology* 218:187–94.

Hauert, C., and M. Doebeli. 2004. "Spatial structure often inhibits the evolution of cooperation in the snowdrift game." *Nature* 428:643–46.

Hauert, C., N. Haiden, and K. Sigmund. 2004. "The dynamics of public goods." *Discrete and Continuous Dynamical Systems B* 4:575–87.

Hauert, C., and G. Szabó. 2003. "Prisoner's Dilemma and public goods games in different geometries: Compulsory versus voluntary interactions." *Complexity* 8:31–38.

Hauert, C., A. Traulsen, H. Brandt, M. A. Nowak, and K. Sigmund. 2007. "Via freedom to coercion: The emergence of costly punishment." *Science* 316:1905–7.

Hauser, M. D. 2006. *Moral Minds: The Nature of Right and Wrong.* New York: Harper Collins.

Heckathorn, D. D. 1996. "The dynamics and dilemmas of collective action." *American Sociological Review* 61:250–77.

Helbing, D. 1992. "A mathematical model for behavioral changes by pair interactions." In *Economic Evolution and Demographic Change: Formal Models in Social Sciences, Lecture Notes in Economics and Mathematical Systems*, vol. 395, eds. G. Haag, U. Mueller, and K. G. Troitzsch, 330–48. Berlin: Springer.

Henrich, J., and R. Boyd. 2001. "Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas." *Journal of Theoretical Biology* 208:79–89.

Henrich, J., R. McElreath, A. Barr, J. Ensminger, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker. 2006. "Costly punishment across human societies." *Science* 312:1767–70.

Herrmann, B., C. Thöni, and S. Gächter. 2008. "Antisocial punishment across societies." *Science* 319:1362–67.

Hilbe, C. 2008. "Memory-one strategien für iterierte spiele mit fehlern." Diploma thesis, Vienna.

Hofbauer, J. 2000. "From Nash and Brown to Maynard Smith: Equilibria, dynamics and ESS." *Selection* 1:81–88.

Hofbauer, J., P. Schuster, and K. Sigmund. 1979. "A note on evolutionary stable strategies and game dynamics." *Journal of Theoretical Biology* 81:609–12.

Hofbauer, J., and K. Sigmund. 1990. "Adaptive dynamics and evolutionary stability." *Applied Mathematics Letters* 3:75–79.

Hofbauer, J., and K. Sigmund. 1998. *Evolutionary Games and Population Dynamics.* Cambridge: Cambridge University Press.

Hofbauer, J., and K. Sigmund. 2003. "Evolutionary game dynamics." *Bulletin of the American Mathematical Society* 40:479–519.

Hutson, V. C. L., and G. T. Vickers. 1995. "The spatial struggle of tit-for-tat and defect." *Philosophical Transactions of the Royal Society B* 348:393–404.

Imhof, L. A., D. Fudenberg, and M. A. Nowak. 2007. "Tit-for-tat or win-stay, lose-shift?" *Journal of Theoretical Biology* 247:574–80.

Isaac, R. M., and J. M. Walker. 1988. "Communication and free-riding behavior: the voluntary contribution mechanism." *Economic Inquiry* 26:585–608.

Johnson, D., and J. Bering. 2006. "Hand of God, mind of man: Punishment and cognition in the evolution of cooperation." *Evolutionary Psychology* 4:219–33.

Johnson, D., P. Stopka, and J. Bell. 2002. "Individual variation evades the Prisoner's Dilemma." *BMC Evolutionary Biology* 2:15.

Kagel, J. H., and A. E. Roth, eds. 1995. *The Handbook of Experimental Economics.* Princeton: Princeton University Press.

Kandori, M. 1992. "Social norms and community enforcement." *The Review of Economic Studies* 59:63–80.

Kandori, M., G. J. Mailath, and R. Rob. 1993. "Learning, mutation, and long run equilibria in games." *Econometrica* 61:29–56.

Kaniovski, Y. M., and H. Young. 1995. "Learning dynamics in games with stochastic perturbations." *Games and Economic Behavior* 11:330–63.

Kendall, G., X. Yao, and S. Y. Chong, eds. 2007. *The Iterated Prisoners' Dilemma: 20 Years On, Advances in Natural Computation*, vol. 4. Singapore: World Scientific.

Keser, C. 2003. "Experimental games for the design of reputation management systems." *IBM Systems Journal* 42:498–506.

Kiers, E. T., R. A. Rousseau, S. A. West, and R. F. Denison. 2003. "Host sanctions and the legume-rhizobium mutualism." *Nature* 425:78–81.

Killingback, T., J. Bieri, and T. Flatt. 2006. "Evolution in group-structured populations can resolve the tragedy of the commons." *Proceedings of the Royal Society B* 273:1477–81.

Killingback, T., and M. Doebeli. 2002. "The continuous Prisoner's Dilemma and the evolution of cooperation through reciprocal altruism with variable investment." *The American Naturalist* 160:421–38.

Kiyonari, T., E. Shimoma, and T. Yamagishi. 2004. "Second-order punishment in one-shot social dilemma." *International Journal of Psychology* 39:329.

Kollock, P. 1993. "'An eye for an eye leaves everyone blind': Cooperation and accounting systems." *American Sociological Review* 58:768–86.

Kollock, P. 1998. "Social dilemmas: The anatomy of cooperation." *Annual Review of Sociology* 24:183–214.

Kraines, D., and V. Kraines. 1989. "Pavlov and the Prisoner's Dilemma." *Theory and Decision* 26:47–79.

Kreps, D. M., and R. Wilson. 1982. "Reputation and imperfect information." *Journal of Economic Theory* 27:253–79.

Kurzban, R., P. DeScioli, and E. O'Brien. 2007. "Audience effects on moralistic punishment." *Evolution and Human Behavior* 28:75–84.

Kurzban, R., and D. Houser. 2005. "Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations." *Proceedings of the National Academy of Sciences* 102:1803–7.

Lehmann, L., and L. Keller. 2006. "The evolution of cooperation and altruism—a general framework and a classification of models." *Journal of Evolutionary Biology* 19:1365–76.

Lehmann, L., F. Rousset, D. Roze, and L. Keller. 2007. "Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment." *The American Naturalist* 170:21–36.

Leimar, O. 1997. "Repeated games: A state space approach." *Journal of Theoretical Biology* 184:471–98.

Leimar, O., and P. Hammerstein. 2001. "Evolution of cooperation through indirect reciprocity." *Proceedings of the Royal Society B* 268:745–53.

Levin, S. A. 1999. *Fragile Dominion: Complexity and the Commons.* Cambridge, MA: Perseus.

Lieberman, E., C. Hauert, and M. A. Nowak. 2005. "Evolutionary dynamics on graphs." *Nature* 433:312–16.

Lindgren, K. 1991. "Evolutionary phenomena in simple dynamics." In *Artificial Life II, Santa Fe Institute studies in the sciences of complexity*, vol. 10, eds. C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, 295–312. Redwood City, CA: Addison-Wesley.

Lindgren, K., and M. G. Nordahl. 1994. "Evolutionary dynamics of spatial games." *Physica D: Nonlinear Phenomena* 75:292–309.

Lotem, A., M. A. Fishman, and L. Stone. 1999. "Evolution of cooperation between individuals." *Nature* 400:226–27.

Lotem, A., M. A. Fishman, and L. Stone. 2003. "From reciprocity to unconditional altruism through signalling benefits." *Proceedings of the Royal Society B* 270:199–205.

Luce, R. D., and H. Raiffa. 1957. *Games and Decisions: Introduction and Critical Survey.*

New York: Wiley.

Masclet, D., C. Noussair, S. Tucker, and M.-C. Villeval. 2003. "Monetary and nonmonetary punishment in the voluntary contributions mechanism." *American Economic Review* 93:366–80.

Masuda, N., and H. Ohtsuki. 2007. "Tag-based indirect reciprocity by incomplete social information." *Proceedings of the Royal Society B* 274:689–95.

Matsui, A. 1992. "Best response dynamics and socially stable strategies." *Journal of Economic Theory* 57:343–62.

May, R. M. 1987. "More evolution of cooperation." *Nature* 327:15–17.

Maynard Smith, J. 1982. *Evolution and the Theory of Games.* Cambridge: Cambridge University Press.

McCabe, K. A., M. L. Rigdon, and V. L. Smith. 2003. "Positive reciprocity and intentions in trust games." *Journal of Economic Behavior & Organization* 52:267–75.

McNamara, J. M., Z. Barta, and A. I. Houston. 2004. "Variation in behaviour promotes cooperation in the Prisoner's Dilemma game." *Nature* 428:745–48.

Milinski, M. 1987. "TIT FOR TAT in sticklebacks and the evolution of cooperation." *Nature* 325:433–35.

Milinski, M., D. Semmann, T. C. M. Bakker, and H.-J. Krambeck. 2001. "Cooperation through indirect reciprocity: image scoring or standing strategy?" *Proceedings of the Royal Society B* 268:2495–501.

Milinski, M., D. Semmann, and H.-J. Krambeck. 2002a. "Donors to charity gain in both indirect reciprocity and political reputation." *Proceedings of the Royal Society B* 269:881–83.

Milinski, M., D. Semmann, and H.-J. Krambeck. 2002b. "Reputation helps solve the 'tragedy of the commons'." *Nature* 415:424–26.

Milinski, M., and C. Wedekind. 1998. "Working memory constrains human cooperation in the Prisoner's Dilemma." *Proceedings of the National Academy of Sciences* 95:13755–58.

Mohtashemi, M., and L. Mui. 2003. "Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altruism." *Journal of Theoretical Biology* 223:523–31.

Molander, P. 1985. "The optimal level of generosity in a selfish, uncertain environment." *The Journal of Conflict Resolution* 29:611–18.

Nakamaru, M., and Y. Iwasa. 2005. "The evolution of altruism by costly punishment in lattice-structured populations: Score-dependent viability versus score-dependent fertility." *Evolutionary Ecology Research* 7:853–70.

Nakamaru, M., and Y. Iwasa. 2006. "The coevolution of altruism and punishment: Role of the selfish punisher." *Journal of Theoretical Biology* 240:475–88.

Neill, D. B. 2001. "Optimality under noise: Higher memory strategies for the alternating Prisoner's Dilemma." *Journal of Theoretical Biology* 211:159–80.

Nowak, M. A. 2006a. *Evolutionary Dynamics: Exploring the Equations of Life.* Cambridge, MA: Harvard University Press.

Nowak, M. A. 2006b. "Five rules for the evolution of cooperation." *Science* 314:1560–63.

Nowak, M. A., S. Bonhoeffer, and R. M. May. 1994. "Spatial games and the maintenance of cooperation." *Proceedings of the National Academy of Sciences* 91:4877–81.

Nowak, M. A., and R. M. May. 1992. "Evolutionary games and spatial chaos." *Nature* 359:826–29.

Nowak, M. A., K. M. Page, and K. Sigmund. 2000. "Fairness versus reason in the Ultimatum game." *Science* 289:1773–75.

Nowak, M. A., and S. Roch. 2007. "Upstream reciprocity and the evolution of gratitude." *Proceedings of the Royal Society B* 274:605–9.

Nowak, M. A., A. Sasaki, C. Taylor, and D. Fudenberg. 2004. "Emergence of cooperation and evolutionary stability in finite populations." *Nature* 428:646–50.

Nowak, M. A., and K. Sigmund. 1989. "Game-dynamical aspects of the Prisoner's Dilemma." *Applied Mathematics and Computation* 30:191–213.

Nowak, M. A., and K. Sigmund. 1990. "The evolution of stochastic strategies in the Prisoner's Dilemma." *Acta Applicandae Mathematicae* 20:247–65.

Nowak, M. A., and K. Sigmund. 1992. "Tit for tat in heterogeneous populations." *Nature* 355:250–53.

Nowak, M. A., and K. Sigmund. 1993. "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game." *Nature* 364:56–58.

Nowak, M. A., and K. Sigmund. 1994. "The alternating Prisoner's Dilemma." *Journal of Theoretical Biology* 168:219–26.

Nowak, M. A., and K. Sigmund. 1998a. "Evolution of indirect reciprocity by image scoring." *Nature* 393:573–77.

Nowak, M. A., and K. Sigmund. 1998b. "The dynamics of indirect reciprocity." *Journal of Theoretical Biology* 194:561–74.

Nowak, M. A., and K. Sigmund. 2005. "Evolution of indirect reciprocity." *Nature* 437:1291–98.

Nowak, M. A., K. Sigmund, and E. El-Sedy. 1995. "Automata, repeated games and noise." *Journal of Mathematical Biology* 33:703–22.

O'Gorman, R., D. S. Wilson, and R. R. Miller. 2005. "Altruistic punishing and helping differ in sensitivity to relatedness, friendship, and future interactions." *Evolution and Human Behavior* 26:375–87.

Ohtsuki, H. 2004. "Reactive strategies in indirect reciprocity." *Journal of Theoretical Biology* 227:299–314.

Ohtsuki, H., and Y. Iwasa. 2004. "How should we define goodness?—reputation dynamics in indirect reciprocity." *Journal of Theoretical Biology* 231:107–20.

Ohtsuki, H., and Y. Iwasa. 2006. "The leading eight: Social norms that can maintain cooperation by indirect reciprocity." *Journal of Theoretical Biology* 239:435–44.

Ohtsuki, H., and Y. Iwasa. 2007. "Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation." *Journal of Theoretical Biology* 244:518–31.

Okuno-Fujiwara, M., and A. Postlewaite. 1995. "Social norms and random matching games." *Games and Economic Behavior* 9:79–109.

Olson, M. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups.* Cambridge, MA: Harvard University Press.

Orbell, J. M., and R. M. Dawes. 1993. "Social welfare, cooperators' advantage, and the option of not playing the game." *American Sociological Review* 58:787–800.

Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* Political Economy of Institutions and Decisions, Cambridge: Cambridge University Press.

Ostrom, E., and J. Walker, eds. 2003. *Trust and Reciprocity: Interdisciplinary Lessons for Experimental Research.* New York: Russell Sage Foundation.

Pacheco, J. M., F. C. Santos, and F. A. C. C. Chalub. 2006. "Stern-judging: A simple, successful norm which promotes cooperation under indirect reciprocity." *PLoS Computational Biology* 2:e178.

Pacheco, J. M., A. Traulsen, H. Ohtsuki, and M. A. Nowak. 2008. "Repeated games and direct reciprocity under active linking." *Journal of Theoretical Biology* 250:723–31.

Page, K. M., and M. A. Nowak. 2001. "A generalized adaptive dynamics framework can describe the evolutionary Ultimatum game." *Journal of Theoretical Biology* 209:173–79.

Page, K. M., and M. A. Nowak. 2002. "Unifying evolutionary dynamics." *Journal of Theoretical Biology* 219:93–98.

Page, K. M., M. A. Nowak, and K. Sigmund. 2000. "The spatial Ultimatum game." *Proceedings of the Royal Society B* 267:2177–82.

Panchanathan, K., and R. Boyd. 2003. "A tale of two defectors: the importance of standing for evolution of indirect reciprocity." *Journal of Theoretical Biology* 224:115–26.

Panchanathan, K., and R. Boyd. 2004. "Indirect reciprocity can stabilize cooperation without the second-order free rider problem." *Nature* 432:499–502.

Pepper, J. W. 2000. "Relatedness in trait group models of social evolution." *Journal of Theoretical Biology* 206:355–68.

Pfeiffer, T., C. Rutte, T. Killingback, M. Taborsky, and S. Bonhoeffer. 2005. "Evolution of cooperation by generalized reciprocity." *Proceedings of the Royal Society B* 272:1115–

20.

Pflügl, C. 2007. "Evolution altruistischen verhaltens durch gruppenselektion." Diploma thesis, Vienna.

Pollock, G., and L. A. Dugatkin. 1992. "Reciprocity and the emergence of reputation." *Journal of Theoretical Biology* 159:25–37.

Poundstone, W. 1992. *Prisoner's Dilemma.* New York: Doubleday.

Price, G. R. 1970. "Selection and covariance." *Nature* 227:520–21.

Price, M. E., L. Cosmides, and J. Tooby. 2002. "Punitive sentiment as an anti-free rider psychological device." *Evolution and Human Behavior* 23:203–31.

Rankin, F. W. 2003. "Communication in Ultimatum games." *Economics Letters* 81:267–71.

Rapoport, A., and A. M. Chammah. 1965. *Prisoner's Dilemma: A Study in Conflict and Cooperation.* Ann Arbor: The University of Michigan Press.

Rege, M., and K. Telle. 2004. "The impact of social approval and framing on cooperation in public good situations." *Journal of Public Economics* 88:1625–44.

Richerson, P. J., and R. Boyd. 2005. *Not by Genes Alone: How Culture Transformed Human Evolution.* Chicago: The University of Chicago Press.

Ridley, M. 1997. *The Origins of Virtue.* New York: Viking Penguin.

Rilling, J. K., D. A. Gutman, T. R. Zeh, G. Pagnoni, G. S. Berns, and C. D. Kilts. 2002. "A neural basis for social cooperation." *Neuron* 35:395–405.

Roberts, G. 2008. "Evolution of direct and indirect reciprocity." *Proceedings of the Royal Society B* 275:173–79.

Roberts, G., and T. N. Sherratt. 1998. "Development of cooperative relationships through increasing investment." *Nature* 394:175–79.

Rockenbach, B., and M. Milinski. 2006. "The efficient interaction of indirect reciprocity and costly punishment." *Nature* 444:718–23.

Rosenthal, R. W. 1979. "Sequences of games with varying opponents." *Econometrica* 47:1353–66.

Rousset, F. 2004. *Genetic Structure and Selection in Subdivided Populations, Monographs in Population Biology,* vol. 40. Princeton: Princeton University Press.

Rubinstein, A. 1986. "Finite automata play the repeated Prisoner's Dilemma." *Journal of Economic Theory* 39:83–96.

Rutte, C., and M. Taborsky. 2007. "Generalized reciprocity in rats." *PLoS Biology* 5:e196.

Sachs, J. L., U. G. Mueller, T. P. Wilcox, and J. J. Bull. 2004. "The evolution of cooperation." *The Quarterly Review of Biology* 79:135–60.

Sandholm, W. H. 2010. *Population Games and Evolutionary Dynamics.* Economic Learning and Social Evolution, Cambridge, MA: MIT Press.

Sanfey, A. G., J. K. Rilling, J. A. Aronson, L. E. Nystrom, and J. D. Cohen. 2003. "The

neural basis of economic decision-making in the Ultimatum game." *Science* 300:1755–58.

Santos, F. C., J. M. Pacheco, and T. Lenaerts. 2006. "Evolutionary dynamics of social dilemmas in structured heterogeneous populations." *Proceedings of the National Academy of Sciences* 103:3490–94.

Schelling, T. C. 1971. "Dynamic models of segregation." *The Journal of Mathematical Sociology* 1:143–86.

Schelling, T. C. 1978. *Micromotives and Macrobehavior.* New York: W. W. Norton & Company.

Schlag, K. H. 1998. "Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits." *Journal of Economic Theory* 78:130–56.

Schuster, P., and K. Sigmund. 1983. "Replicator dynamics." *Journal of Theoretical Biology* 100:533–38.

Sefton, M., R. Shupp, and J. M. Walker. 2007. "The effect of rewards and sanctions in provision of public goods." *Economic Inquiry* 45:671–90.

Seinen, I., and A. Schram. 2006. "Social status and group norms: Indirect reciprocity in a repeated helping experiment." *European Economic Review* 50:581–602.

Sell, J., and R. K. Wilson. 1999. "The maintenance of cooperation: Expectations of future interaction and the trigger of group punishment." *Social Forces* 77:1551–71.

Selten, R., and P. Hammerstein. 1984. "Gaps in Harley's argument on evolutionarily stable learning rules and in the logic of 'tit for tat'." *Behavioral and Brain Sciences* 7:115–16.

Semmann, D., H.-J. Krambeck, and M. Milinski. 2003. "Volunteering leads to rock-paper-scissors dynamics in a public goods game." *Nature* 425:390–93.

Semmann, D., H.-J. Krambeck, and M. Milinski. 2004. "Strategic investment in reputation." *Behavioral Ecology and Sociobiology* 56:248–52.

Sherratt, T. N., and G. Roberts. 2001. "The importance of phenotypic defectors in stabilizing reciprocal altruism." *Behavioral Ecology* 12:313–17.

Sigmund, K. 1995. *Games of Life: Explorations in Ecology, Evolution, and Behaviour.* Harmondsworth: Penguin Press.

Sigmund, K. 2007. "Punish or perish? Retaliation and collaboration among humans." *Trends in Ecology & Evolution* 22:593–600.

Sigmund, K., C. Hauert, and M. A. Nowak. 2001. "Reward and punishment." *Proceedings of the National Academy of Sciences* 98:10757–62.

Silk, J. B. 2006. "Who are more helpful, humans or chimpanzees?" *Science* 311:1248–49.

Skyrms, B. 2004. *The Stag Hunt and the Evolution of Social Structure.* Cambridge: Cambridge University Press.

Skyrms, B., and R. Pemantle. 2000. "A dynamic model of social network formation."

Proceedings of the National Academy of Sciences 97:9340–46.

Small, D. A., and G. Loewenstein. 2005. "The devil you know: The effects of identifiability on punishment." *Journal of Behavioral Decision Making* 18:311–18.

Sober, E., and D. S. Wilson. 1998. *Unto Others: the Evolution and Psychology of Unselfish Behavior.* Cambridge, MA: Harvard University Press.

Sommerfeld, R. D., H.-J. Krambeck, D. Semmann, and M. Milinski. 2007. "Gossip as an alternative for direct observation in games of indirect reciprocity." *Proceedings of the National Academy of Sciences* 104:17435–40.

Stephens, D. W., C. M. McLinn, and J. R. Stevens. 2002. "Discounting and reciprocity in an iterated prisoner's dilemma." *Science* 298:2216–18.

Stevens, J. R., F. A. Cushman, and M. D. Hauser. 2005. "Evolving the psychological mechanisms for cooperation." *Annual Review of Ecology, Evolution, and Systematics* 36:499–518.

Sugden, R. 1986. *The Economics of Rights, Co-operation and Welfare.* Oxford: Basil, Blackwell.

Suzuki, S., and E. Akiyama. 2007a. "Evolution of indirect reciprocity in groups of various sizes and comparison with direct reciprocity." *Journal of Theoretical Biology* 245:539–52.

Suzuki, S., and E. Akiyama. 2007b. "Three-person game facilitates indirect reciprocity under image scoring." *Journal of Theoretical Biology* 249:93–100.

Szabó, G., and G. Fáth. 2007. "Evolutionary games on graphs." *Physics Reports* 446:97–216.

Szabó, G., and C. Hauert. 2002. "Phase transitions and volunteering in spatial public goods games." *Physical Review Letters* 89:118101.

Szabó, G., and J. Vukov. 2004. "Cooperation for volunteering and partially random partnerships." *Physical Review E* 69:036107.

Takahashi, N., and R. Mashima. 2003. "The emergence of indirect reciprocity: Is the standing strategy the answer?" Working paper, Hokkaido University.

Takahashi, N., and R. Mashima. 2006. "The importance of subjectivity in perceptual errors on the emergence of indirect reciprocity." *Journal of Theoretical Biology* 243:418–36.

Taylor, C., D. Fudenberg, A. Sasaki, and M. A. Nowak. 2004. "Evolutionary game dynamics in finite populations." *Bulletin of Mathematical Biology* 66:1621–44.

Taylor, P. D., T. Day, and G. Wild. 2007. "Evolution of cooperation in a finite homogeneous graph." *Nature* 447:469–72.

Taylor, P. D., and S. A. Frank. 1996. "How to make a kin selection model." *Journal of Theoretical Biology* 180:27–37.

Taylor, P. D., and L. B. Jonker. 1978. "Evolutionary stable strategies and game dynamics." *Mathematical Biosciences* 40:145–56.

Traulsen, A., C. Hauert, H. De Silva, M. A. Nowak, and K. Sigmund. 2009. "Exploration dynamics in evolutionary games." *Proceedings of the National Academy of Sciences* 106:709–12.

Trivers, R. 1971. "The evolution of reciprocal altruism." *The Quarterly Review of Biology* 46:35–57.

Trivers, R. 1985. *Social Evolution.* Menlo Park, CA: Benjamin Cummings.

Trivers, R. 2002. *Natural Selection and Social Theory: Selected Papers of Robert Trivers.* New York: Oxford University Press.

Trivers, R. 2006. "Reciprocal altruism: 30 years later." In *Cooperation in Primates and Humans: Mechanisms and Evolution*, eds. P. M. Kappeler and C. P. van Schaik, 67–83. Berlin: Springer.

Vincent, T. L., and J. S. Brown. 2005. *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics.* Cambridge: Cambridge University Press.

Wahl, L. M., and M. A. Nowak. 1999. "The continuous Prisoner's Dilemma: I. Linear reactive strategies." *Journal of Theoretical Biology* 200:307–21.

Walker, J. M., and M. A. Halloran. 2004. "Rewards and sanctions and the provision of public goods in one-shot settings." *Experimental Economics* 7:235–47.

Warneken, F., and M. Tomasello. 2006. "Altruistic helping in human infants and young chimpanzees." *Science* 311:1301–3.

Wedekind, C., and V. A. Braithwaite. 2002. "The long-term benefits of human generosity in indirect reciprocity." *Current Biology* 12:1012–15.

Wedekind, C., and M. Milinski. 2000. "Cooperation through image scoring in humans." *Science* 288:850–52.

Weibull, J. W. 1995. *Evolutionary Game Theory.* Cambridge, MA: MIT Press.

Wenseleers, T., and F. L. W. Ratnieks. 2006. "Comparative analysis of worker reproduction and policing in eusocial Hymenoptera supports relatedness theory." *The American Naturalist* 168:E163–79.

West, S. A., A. S. Griffin, and A. Gardner. 2007. "Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection." *Journal of Evolutionary Biology* 20:415–32.

West, S. A., I. Pen, and A. S. Griffin. 2002. "Cooperation and competition between relatives." *Science* 296:72–75.

Xiao, E., and D. Houser. 2005. "Emotion expression in human punishment behavior." *Proceedings of the National Academy of Sciences* 102:7398–401.

Yamagishi, T. 1986. "The provision of a sanctioning system as a public good." *Journal of Personality and Social Psychology* 51:110–16.

Yamagishi, T., N. Jin, and T. Kiyonari. 1999. "Bounded generalized reciprocity: Ingroup boasting and ingroup favoritism." *Advances in Group Processes* 16:161–97.

Zeeman, E. C. 1980. "Population dynamics from game theory." In *Global Theory of Dynamical Systems*, *Lecture Notes in Mathematics*, vol. 819, eds. Z. Nitecki and C. Robinson, 471–97. Berlin: Springer.

索　引

多疑的 (suspicious), 54, 85, 87

M

N

P

Q

S

T

W

X

Y

严格 (strict), 9, 18, 26-28, 33, 34, 37, 45, 63, 104, 119, 127

Z

其他

《现代数学译丛》已出版书目

（按出版时间排序）

1　椭圆曲线及其在密码学中的应用——导引　2007.12　〔德〕Andreas Enge　著　吴　铤　董军武　王明强　译
2　金融数学引论——从风险管理到期权定价　2008.1　〔美〕Steven Roman　著　邓欣雨　译
3　现代非参数统计　2008.5　〔美〕Larry Wasserman　著　吴喜之　译
4　最优化问题的扰动分析　2008.6　〔法〕J. Frédéric Bonnans　〔美〕Alexander Shapiro　著　张立卫　译
5　统计学完全教程　2008.6　〔美〕Larry Wasserman　著　张　波　等译
6　应用偏微分方程　2008.7　〔英〕John Ockendon, Sam Howison, Andrew Lacey & Alexander Movchan　著　谭永基　程　晋　蔡志杰　译
7　有向图的理论、算法及其应用　2009.1　〔丹〕J. 邦詹森　〔英〕G. 古廷　著　姚兵　张忠辅　译
8　微分方程的对称与积分方法　2009.1〔加〕乔治 W. 布卢曼　斯蒂芬 C. 安科　著　闫振亚　译
9　动力系统入门教程及最新发展概述　2009.8　〔美〕Boris Hasselblatt & Anatole Katok　著　朱玉峻　郑宏文　张金莲　阎欣华　译　胡虎翼　校
10　调和分析基础教程　2009.10　〔德〕Anton Deitmar　著　丁勇　译
11　应用分支理论基础　2009.12　〔俄〕尤里·阿·库兹涅佐夫　著　金成桴　译
12　多尺度计算方法——均匀化及平均化　2010.6　Grigorios A. Pavliotis, Andrew M. Stuart　著　郑健龙　李友云　钱国平　译
13　最优可靠性设计：基础与应用　2011.3〔美〕Way Kuo, V. Rajendra Prasad, Frank A.Tillman, Ching-Lai Hwang　著　郭进利　闫春宁　译　史定华　校
14　非线性最优化基础　2011.4　〔日〕Masao Fukushima　著　林贵华　译
15　图像处理与分析：变分，PDE，小波及随机方法　2011.6　Tony F. Chan, Jianhong (Jackie) Shen　著　陈文斌，程晋　译
16　马氏过程　2011.6　〔日〕福岛正俊　竹田雅好　著　何萍　译　应坚刚　校

17 合作博弈理论模型 2011.7 〔罗〕Rodica Branzei 〔德〕Dinko Dimitrov 〔荷〕Stef Tijs 著 刘小冬 刘九强 译

18 变分分析与广义微分 I：基础理论 2011.9 〔美〕 Boris S. Mordukhovich 著 赵亚莉 王炳武 钱伟懿 译

19 随机微分方程导论应用(第 6 版) 2012.4 〔挪〕Bernt Øksendal 著 刘金山 吴付科 译

20 金融衍生产品的数学模型 2012.4 郭宇权(Yue-Kuen Kwok) 著 张寄洲 边保军 徐承龙 等 译

21 欧拉图与相关专题 2012.4 〔英〕Herbert Fleischner 著 孙志人 李 皓 刘桂真 刘振宏 束金龙 译 张 昭 黄晓晖 审校

22 重分形：理论及应用 2012.5 〔美〕戴维·哈特 著 华南理工分形课题组 译

23 组合最优化：理论与算法 2014. 1 〔德〕 Bernhard Korte Jens Vygen 著 姚恩瑜 林治勋 越民义 张国川 译

24 变分分析与广义微分 II：应用 2014. 1 〔美〕 Boris S. Mordukhovich 著 李 春 王炳武 赵亚莉 王 东 译

25 算子理论的 Banach 代数方法(原书第二版) 2014.3 〔美〕 Ronald G. Douglas 著 颜 军 徐胜芝 舒永录 蒋卫生 郑德超 孙顺华 译

26 Bäcklund 变换和 Darboux 变换——几何与孤立子理论中的应用 2015.5 〔澳〕C. Rogers W. K. Schief 著 周子翔 译

27 凸分析及应用捷径 2015.9 〔美〕 Boris S. Mordukhovich Nguyen Mau Nam 著 赵亚莉 王炳武 译

28 利己主义的数学解析 2017.8 〔奥〕 K. Sigmund 著 徐金亚 杨 静 汪芳 译

(O-6987.31)



科学出版社互联网入口
科学数理分社
电 话：(010) 64011058
Email：chenyuzhuo@mail.sciencep.com
销售分类建议：高等数学

www.sciencep.com



定 价：68.00 元